

Tách phong đối tượng người trong ảnh bằng phân vùng ngữ nghĩa FCN-CRFs

Nguyễn Nhật Tín
(B1507321)

Giảng viên hướng dẫn:
TS. Thái Minh Tuấn

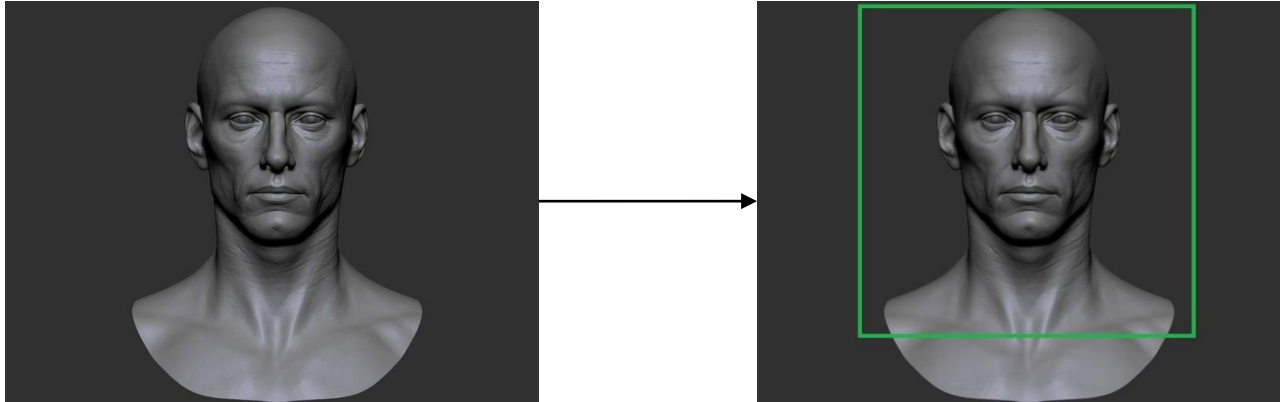
Nội dung

- Giới thiệu
- Cơ sở lý thuyết
- Tập dữ liệu
- Đánh giá
- Kết luận
- Chạy thực nghiệm

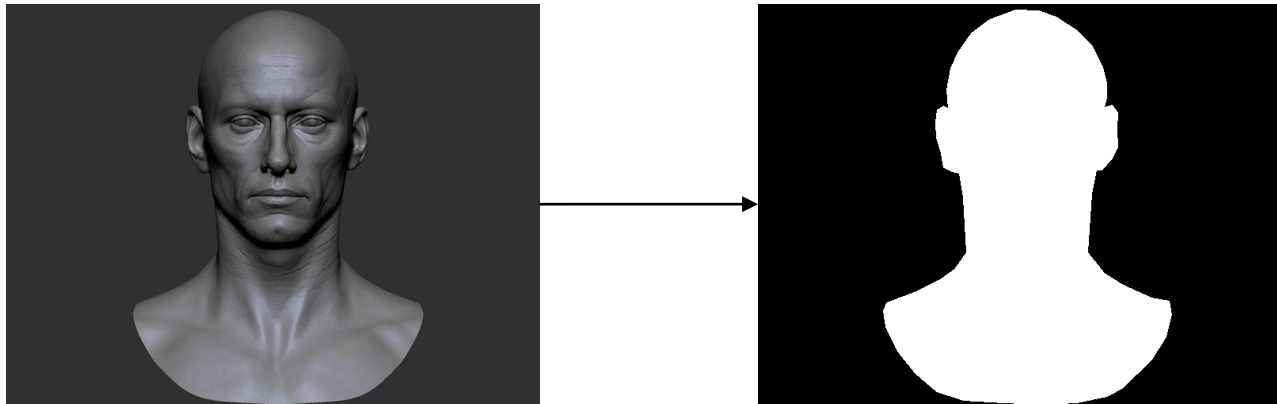
Nội dung

- **Giới thiệu**
- Cơ sở lý thuyết
- Tập dữ liệu
- Đánh giá
- Kết luận
- Chạy thực nghiệm

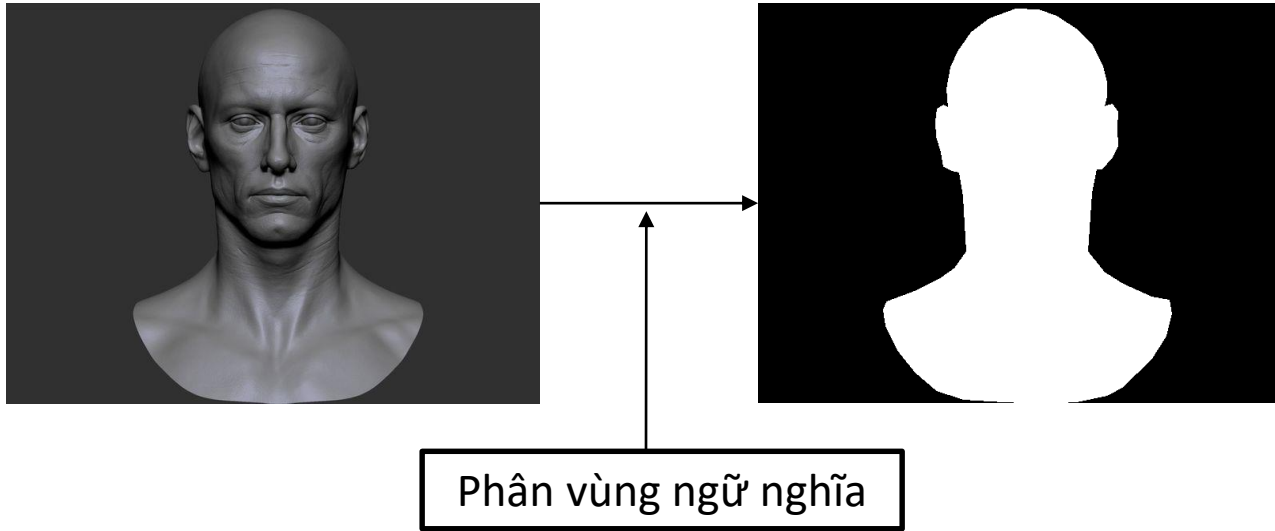
Giới thiệu



Giới thiệu



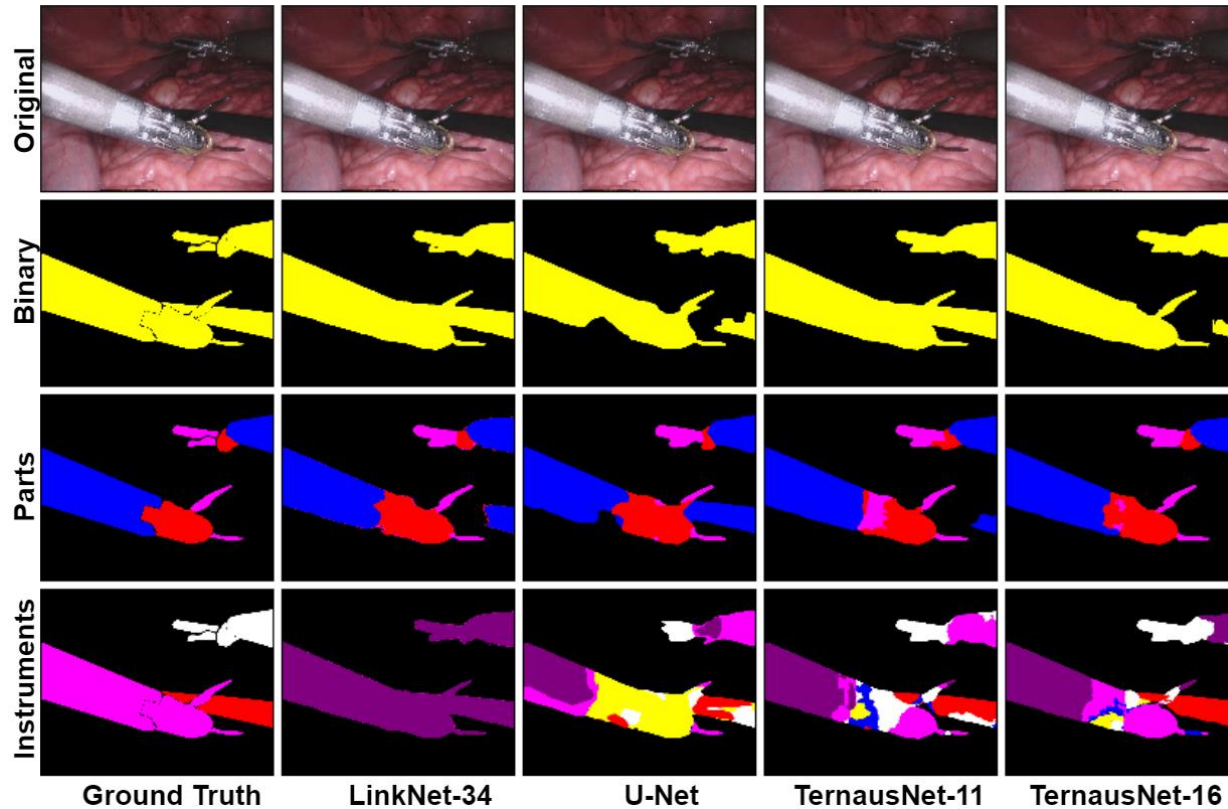
Giới thiệu



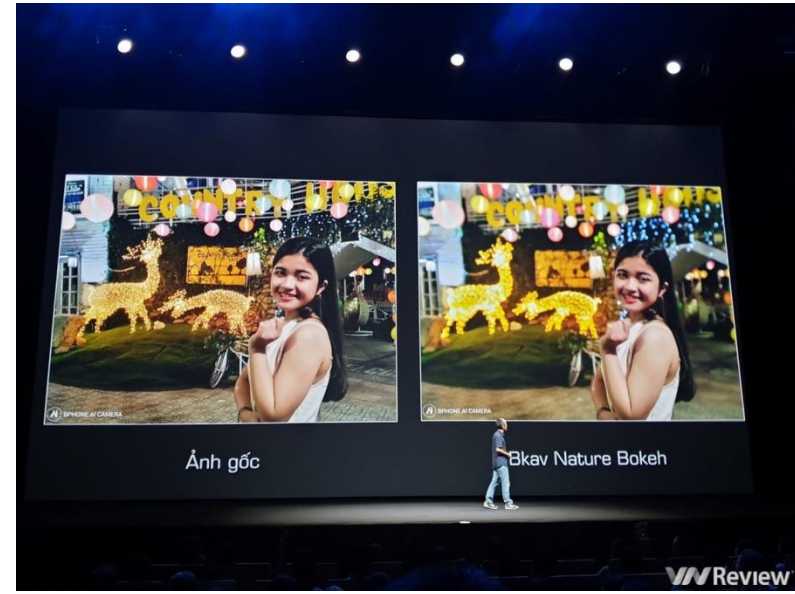
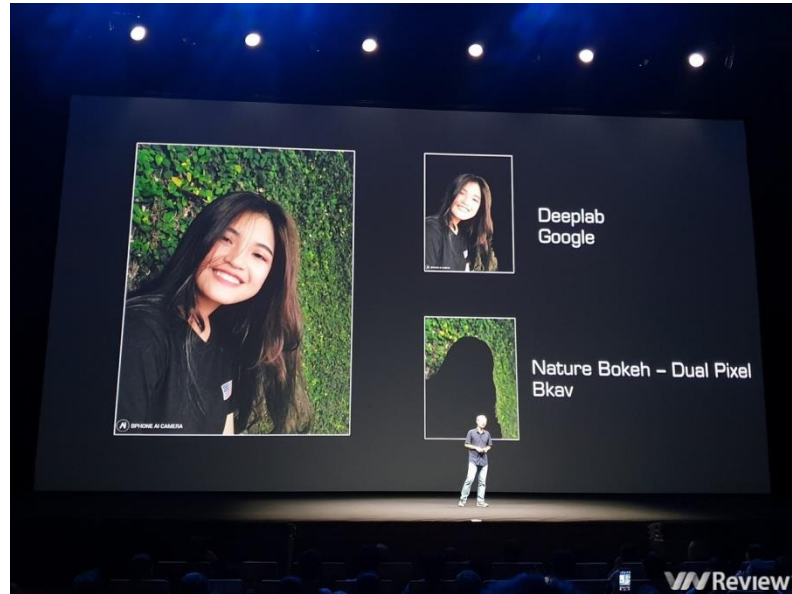
Định vị đối tượng trên đường



Phẫu thuật y tế



Công nghệ máy ảnh điện thoại



Ứng dụng Snow



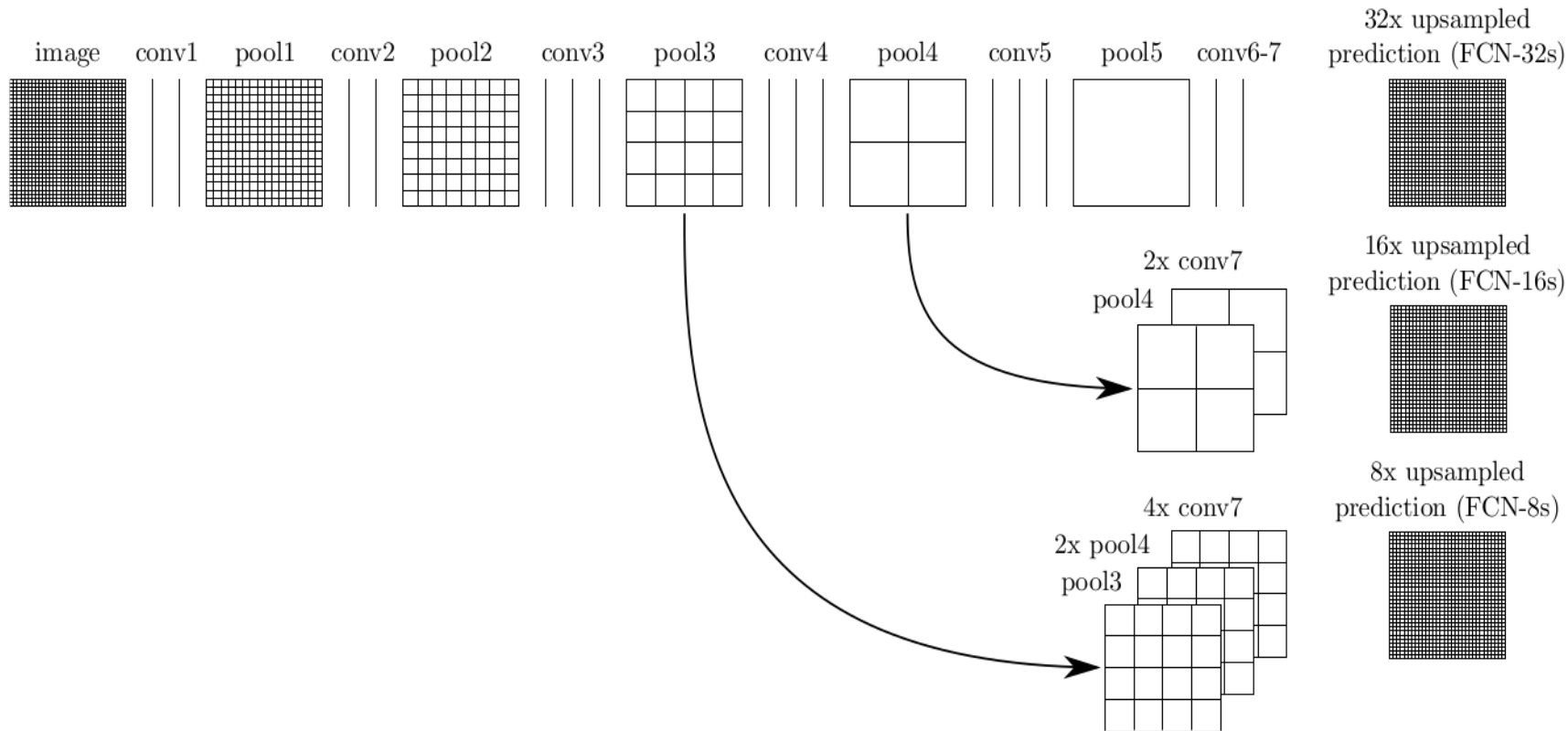
Biến đổi phong cảnh của ảnh



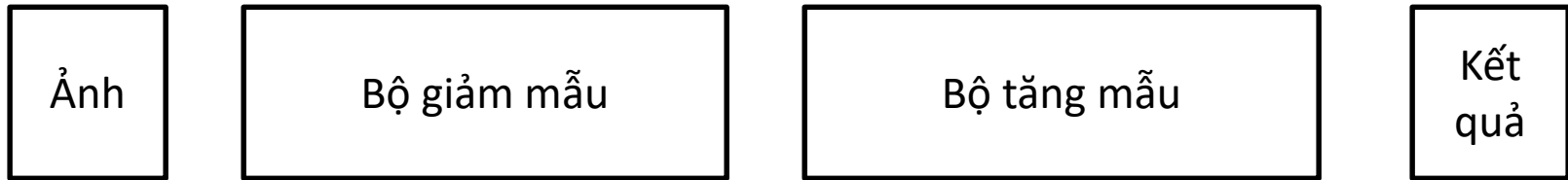
Nội dung

- Giới thiệu
- **Cơ sở lý thuyết**
- Tập dữ liệu
- Đánh giá
- Kết luận
- Chạy thực nghiệm

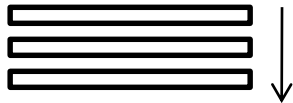
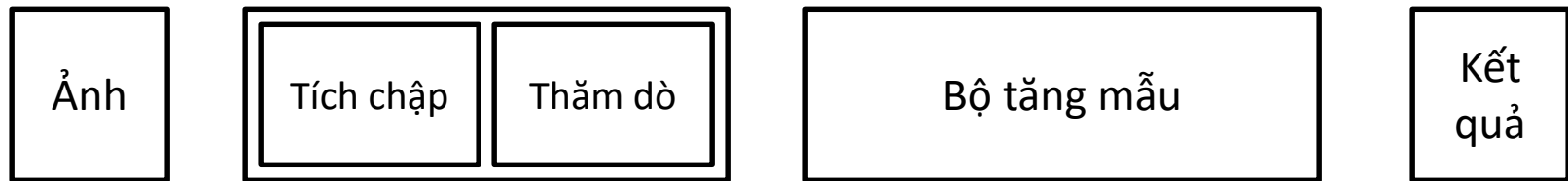
Mạng tích chập đầy đủ (FCN)



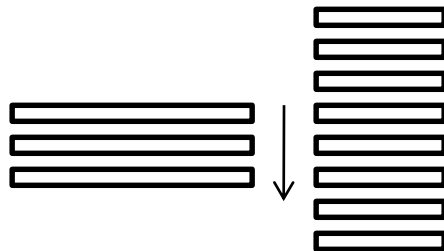
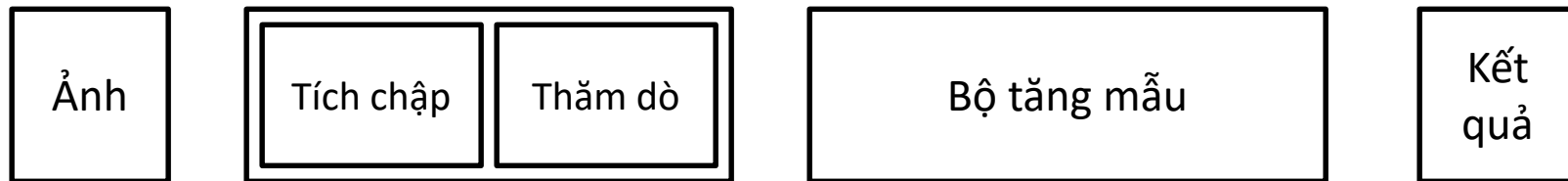
Mạng tích chập đầy đủ (FCN)



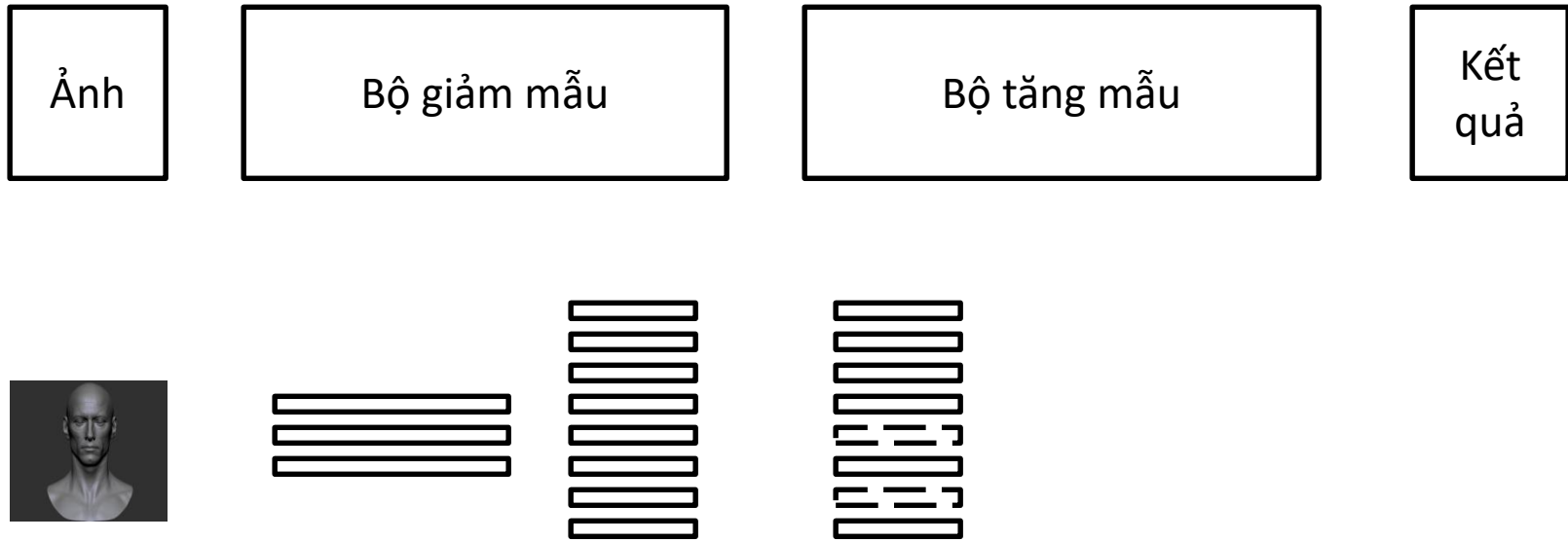
Mạng tích chập đầy đủ (FCN)



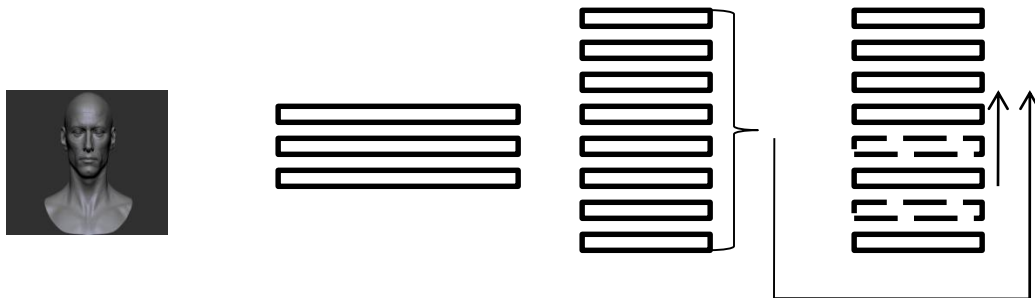
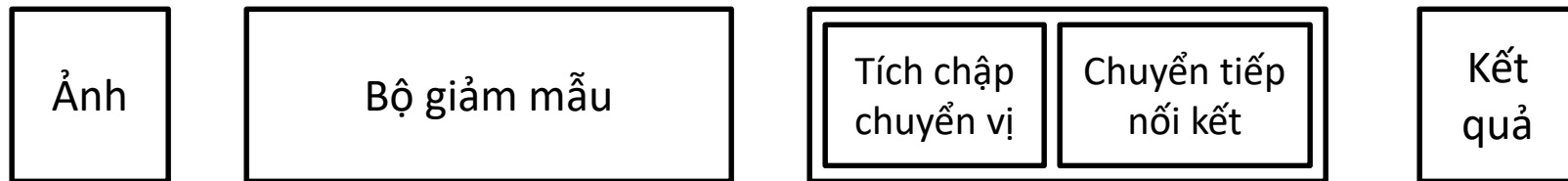
Mạng tích chập đầy đủ (FCN)



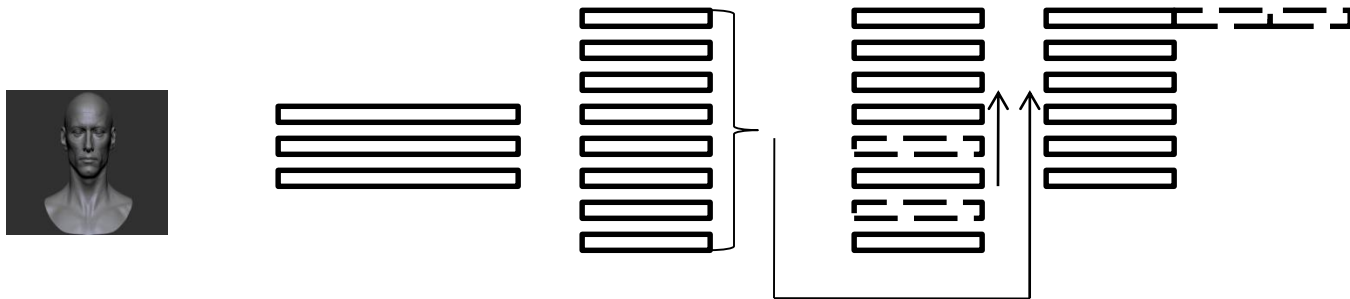
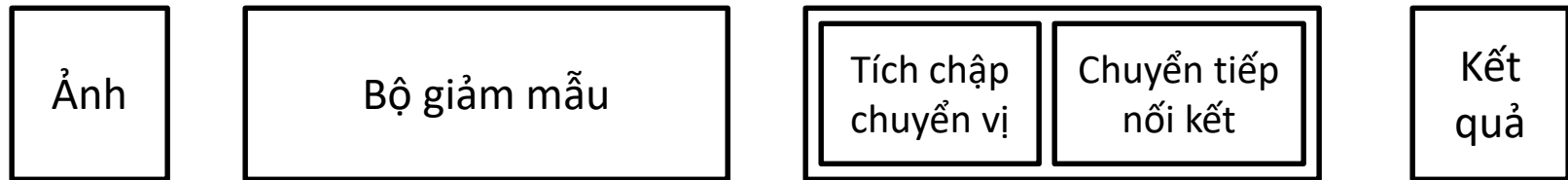
Mạng tích chập đầy đủ (FCN)



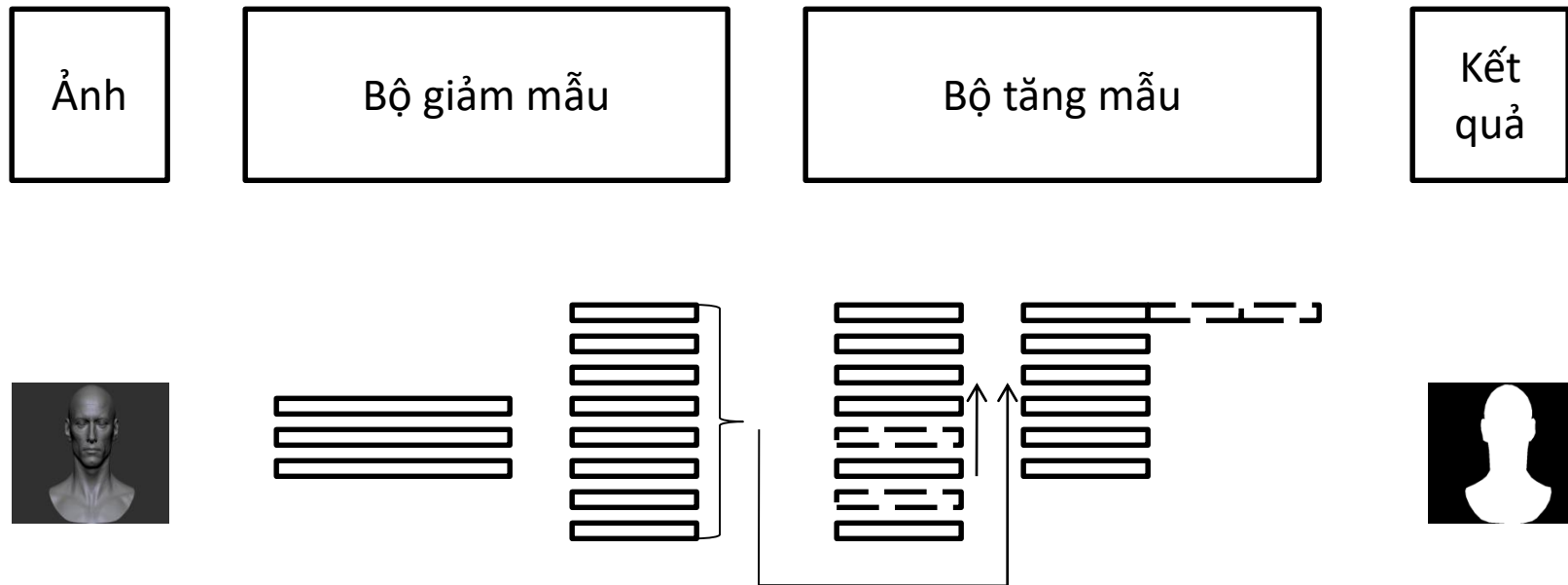
Mạng tích chập đầy đủ (FCN)



Mạng tích chập đầy đủ (FCN)



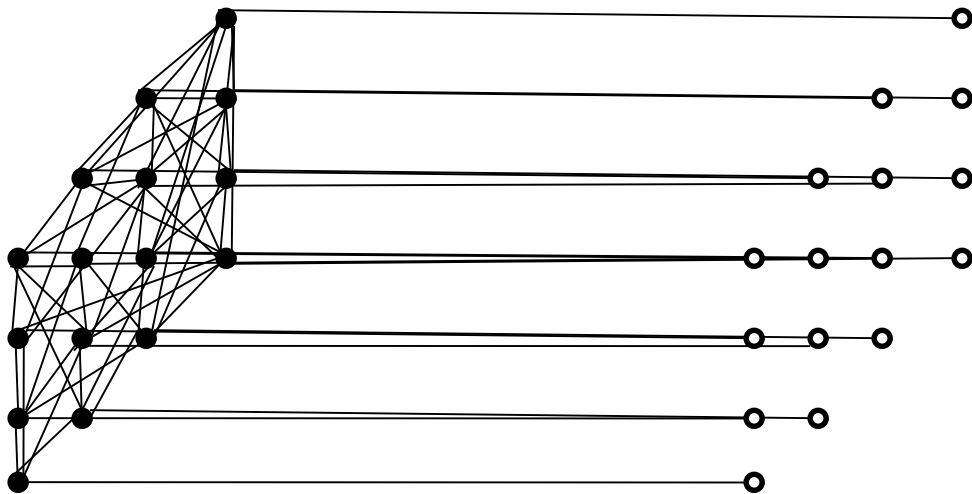
Mạng tích chập đầy đủ (FCN)



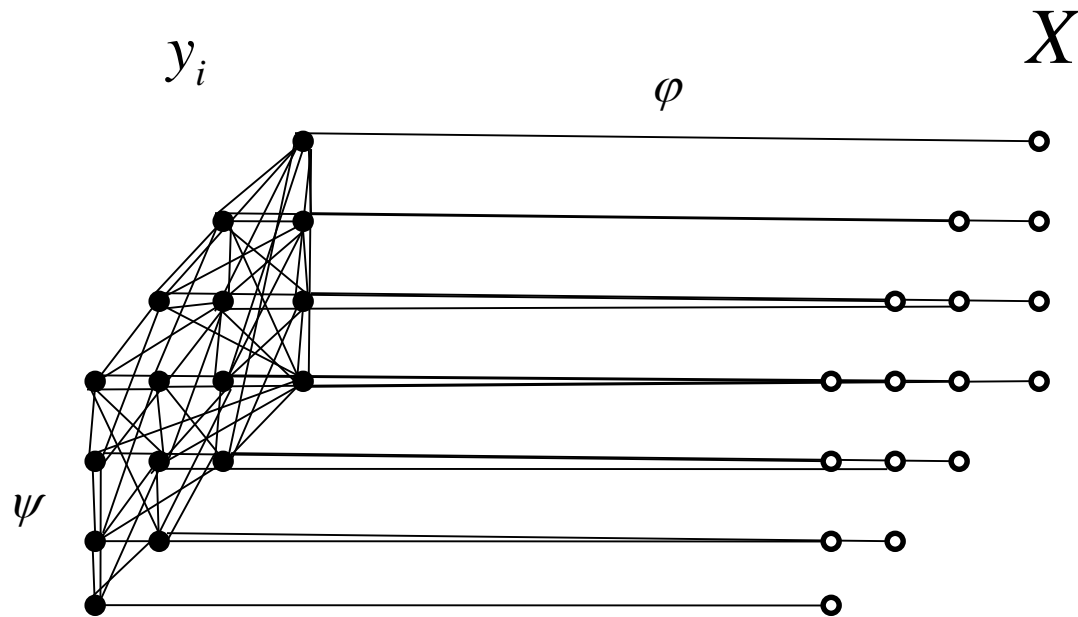
Trường điều kiện ngẫu nhiên (CRFs)



Trường điều kiện ngẫu nhiên (CRFs)

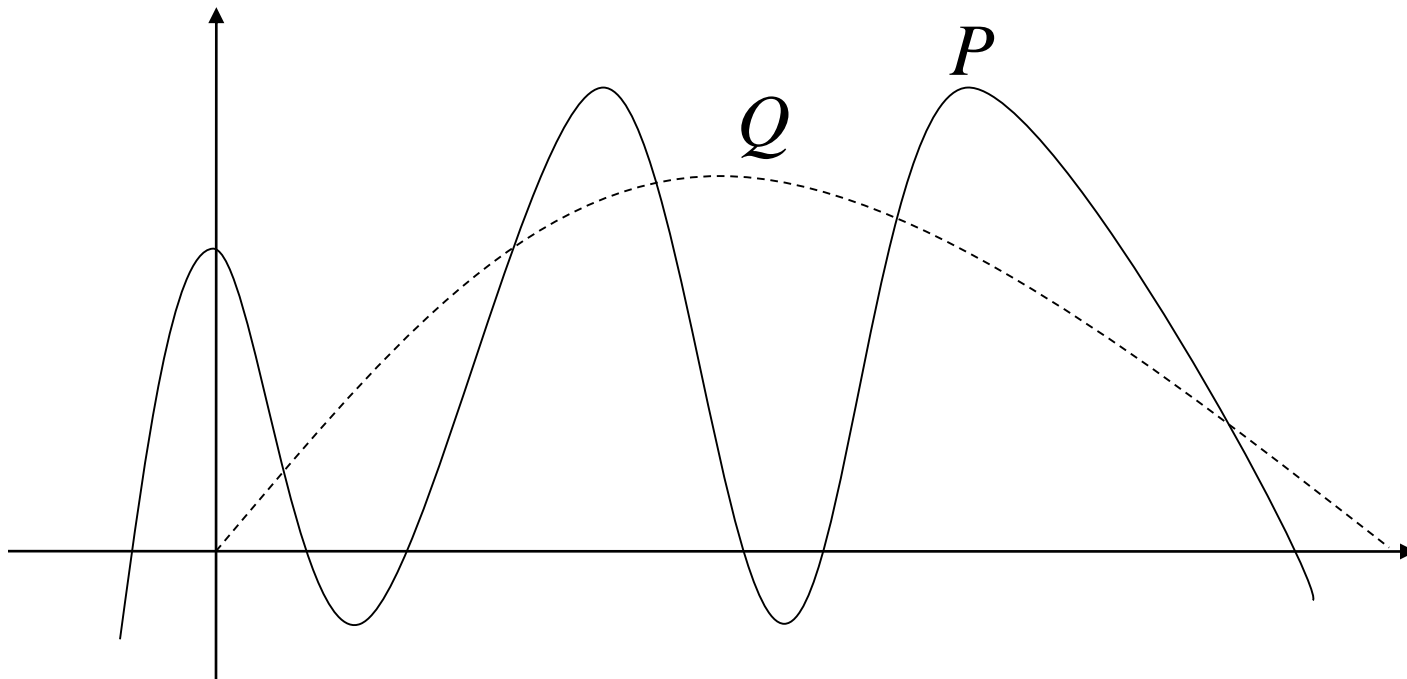


Trường điều kiện ngẫu nhiên (CRFs)

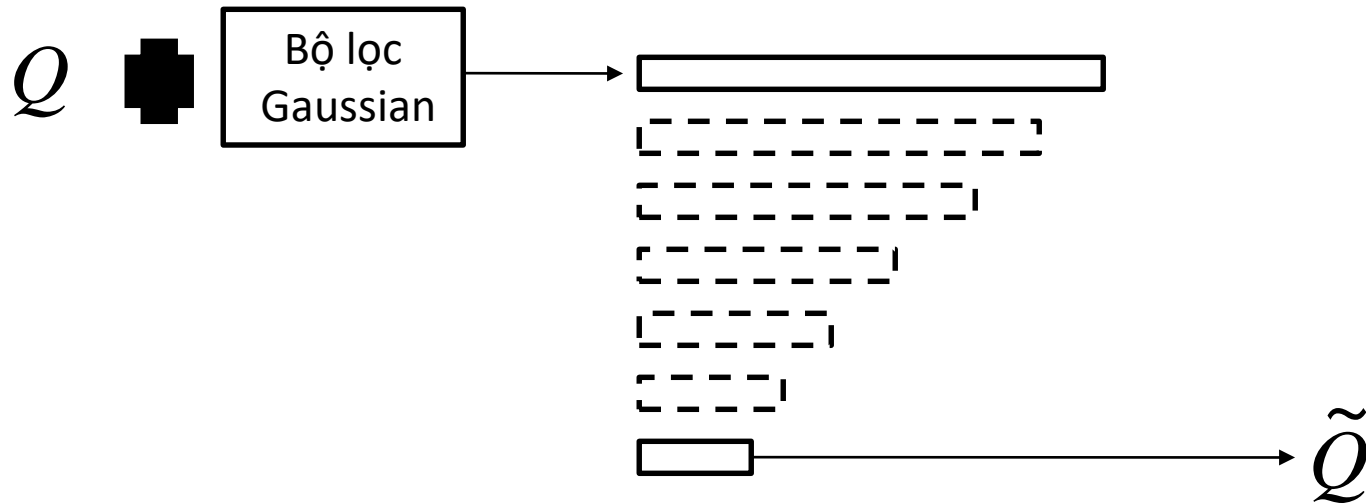


$$E(y | X) = \sum_i \phi(y_i | X) + \sum_{ij} \psi(y_i, y_j) \longrightarrow P$$

Trường điều kiện ngẫu nhiên (CRFs)



Trường điều kiện ngẫu nhiên (CRFs)



Nội dung

- Giới thiệu
- Cơ sở lý thuyết
- **Tập dữ liệu**
- Đánh giá
- Kết luận
- Chạy thực nghiệm

Tập dữ liệu

- Tập hợp 2 tập dữ liệu Face/Headseg và Part Labels
- 525 tấm ảnh (224 x 224) gồm:
 - Tập huấn luyện 395 ảnh
 - Tập kiểm tra 130 ảnh
- Sử dụng 2 loại nhãn [đỏ, lục, lam] gồm:
 - [255,255,255] = trắng = người
 - [0,0,0] = đen = phong nền

Nội dung

- Giới thiệu
- Cơ sở lý thuyết
- Tập dữ liệu
- **Đánh giá**
- Kết luận
- Chạy thực nghiệm

Kiểm tra FCNs

| Tổng = 130 ảnh | F1 (tiền cảnh) | F1 (hậu cảnh) | MeanIU |
|-----------------------|-----------------------|----------------------|---------------|
| FCN-32s | 0.819 | 0.925 | <i>0.777</i> |
| FCN-16s | 0.861 | 0.937 | <i>0.818</i> |
| FCN-8s | 0.863 | 0.939 | <i>0.822</i> |

Tối ưu CRFs

| Tổng = 130 ảnh | F1 (tiền cảnh) | F1 (hậu cảnh) | MeanIU |
|----------------|----------------|---------------|--------------|
| FCN-8s-CRFs | 0.780 | 0.914 | <i>0.741</i> |

Cấu hình phần cứng kiểm thử

- Hệ điều hành Windows 7 64-bit
- CPU Core i5-2540M @ 2.60GHz
- RAM 8GB

Nội dung

- Giới thiệu
- Cơ sở lý thuyết
- Tập dữ liệu
- Đánh giá
- **Kết luận**
- Chạy thực nghiệm

Giới hạn

- Không thể thực hiện trên video và thời gian thực
- Kết quả tối ưu tồn tại sự không liên tục
- Không thể thu được kết quả tốt khi phân vùng ngữ nghĩa với các đối tượng là nhiều người và các loại vật thể khác nhau

Hướng phát triển

- Sử dụng các loại giải thuật phức tạp hơn như SegNet, UNet, DeepLab thay cho FCNs
- Cải tiến trường điều kiện ngẫu nhiên cải thiện sự mất cân bằng các lớp
- Bổ sung hình ảnh phong phú cả về không gian, thời gian và vật thể trong ảnh

Nội dung

- Giới thiệu
- Cơ sở lý thuyết
- Tập dữ liệu
- Đánh giá
- Kết luận
- **Chạy thực nghiệm**

Chạy thực nghiệm

Nhận xét và bổ sung

Xin chân thành cảm ơn