**BÁO CÁO BASIC AI/ML TUẦN 8+9**

Yêu cầu hoàn thành các bài tập về AI/ML chương 6.

Bài 1: Hãy vẽ sơ đồ khối của giải pháp để nhận diện đối tượng mang giới tính nam, mặc quẩn đùi đen và áo ba lỗ đỏ.

A diagram of a black grid

Description automatically generated with medium confidence

Bài 2: Thực hiện phép deconvolution sau, biết rằng:

Output feature map =

Kernel =

TH 1: Stride = 1

TH 2: Stride = 2

*Ta có :O = ((W - K + 2P) / S) + 1* với W là kích thước đầu vào, K là kích thước Kernel, P là padding, S là Stride.

TH 1: Stride = 1, K=2, Padding =0 🡪 W=3. Ma trận cần tìm có kích thước 3x3.

Đặt ma trận ban đầu . Ta có các ma trận cần tìm phải thỏa mãn hệ phương trình sau: . Ví dụ : là 1 ma trận thỏa mãn.

TH 2: Stride = 2, K=2, Padding =0 🡪 W=4. Ma trận cần tìm có kích thước 4x4.

Đặt ma trận ban đầu là , ta có được hệ phương trình như sau

Như vậy, tập các ma trận cần tìm phải thỏa mãn hệ phương trình trên.

Câu 4: Thực hiện phép atrous convolution sau, biết rằng:

Input feature map =

Kernel =

Stride = 1

Rate = 2

Padding = 0

Bài làm được hoàn thiện trong bài 3.ipynb

Câu 5: Khi thực hiện phép tích chập giãn nở ở câu 4, số lần kernel trượt bằng bao nhiêu?

Khi thực hiện phép tích chập ở câu trên, số lần kernel trượt bằng số phần tử đầu ra -1 🡪 9-1=8 lần trượt.

Câu 6: Có bao nhiêu phép nhân FP đc thực hiện khi thực hiện phép toán tích chập ở câu 4.

Phép nhân floating point: phép nhân giũa 2 số dấu phẩy động, đây là phép toán căn bản khi thực hiện tích chập giũa 2 ma trận.

Số thành phần của ma trận đầu ra là 9

Với mỗi thành phần, ta cần thực hiện 25 phép nhân ứng với mỗi thành phần tương ứng của ma trận kernel giãn và ma trận đầu vào

* 9 điểm trong output feature map x 25 phép nhân FP mỗi điểm = 225 phép nhân FP

Vậy, số lượng phép nhân FP được thực hiện khi thực hiện phép tích chập ở câu 4 là 225.

Câu 7: Hãy liệt kê 4 giải pháp khác nhau cho bài toán định danh khuôn mặt (face recognition)

4 phương pháp khác nhau để giải quyết bài toán nhận diện khuôn mặt:

* Knowledge – based:
* Đối với kỹ thuật dựa trên kiến thức để nhận diện khuôn mặt, một tập hợp các quy tắc phải được tuân theo, và nó phụ thuộc vào sự hiểu biết của con người. Ví dụ, một khuôn mặt phải có mũi, mắt và miệng với khoảng cách và vị trí nhất định so với nhau.
* Nhược điểm lớn nhất là khó khăn trong việc phát triển một tập hợp các tiêu chí phù hợp. Nếu các tiêu chí quá mơ hồ hoặc quá cụ thể, có thể xảy ra nhiều trường hợp dương tính giả (Ví dụ như mặt nạ có đầy đủ các bộ phận và khoảng cách phù hợp sẽ dễ gây ra False Postive)
* Phương pháp này một mình không đủ và không thể xác định nhiều khuôn mặt trong một số lượng lớn hình ảnh.
* Feature – based:
* Kỹ thuật dựa trên đặc trưng được sử dụng để tìm khuôn mặt bằng cách trích xuất các đặc điểm cấu trúc từ các khuôn mặt đang được tìm kiếm. Nó được đào tạo sau đó sử dụng để phân biệt giữa các phần khuôn mặt và các phần không phải khuôn mặt của cơ thể.
* Đây được coi là phương pháp hiệu quả trong việc nhận diện nhiều khuôn mặt, tuy nhiên cần 1 lượng lớn dữ liệu để đào tạo. Ngoài ra, phương pháp trên có thể gặp khó khăn khi mặt bị biến dạng bởi các yếu tố như góc nhìn, môi trường, biểu cảm khuôn mặt.
* Matching Templates:
* Về cơ bản phương pháp này sử dụng các mẫu khuôn mặt đã được xác định trước hoặc tham số hóa để xác định hoặc phát hiện khuôn mặt bằng cách so sánh chúng với các bức ảnh đầu vào.
* Mặc dù phương pháp này đơn giản để phát triển, nhưng nó không đủ cho việc phát hiện khuôn mặt nếu khuôn mặt đầu vào không có trong tập dữ liệu có sẵn. Phương pháp này sử dụng để tìm kiếm các khớp giữa hai mẫu khuôn mặt.
* Appearance – based:
* Phương pháp này sử dụng một bộ dữ liệu lớn các hình ảnh khuôn mặt để đào tạo mô hình nhận diện, thường sử dụng các kỹ thuật học máy và phân tích thống kê.
* Đây là phương pháp có toàn diện với hiệu suất cao và có khả năng nhận diện chính xác trong nhiều điều kiện khác nhau. Tuy nhiên cần nhiều dữ liệu và tài nguyên tính toán để đào tạo mô hình.

Câu 8: Trong mô hình yolov4, nếu tăng số anchor box từ 9 lên 15, hỏi có bao nhiêu predicted bounding box được tạo ra tại đầu ra của mô hình Yolov4 khi train trên tập VOC

Do các anchor box được tạo ra , gán vào các feature map và sau đó lựa chon anchor box có IOU lớn nhất 🡪 số lượng predicted bouding box tạo ra tương tự trước khin thay đổi số lượng anchor box.

Câu 9: Cho Ground Truth Mask và Predicted Mask như sau, hãy tính accuracy của segmentation model này.

Ground truth mask =

Predicted mask =