nếu như nhận diện mèo là input ảnh predict có phải mèo hay không

giờ với computervision thì ta cần phải vẽ đc box của vật trong một bức ảnh

ngoài ra kích thước ảnh là không cố định, có thể rất lớn, rất nhỏ

rất lớn sẽ khiến input có quá nhiều pixel dẫn tới số param tăng theo

VD: 3tr điểm ảnh, 1000 nút nhân vs nhau = 3 tỷ => hiệu năng kém

Edge Detection

Vertical Edege Detector

Diagram

Description automatically generated

ma trận tìm ra cạnh học là ở trên

việc nhân như vậy ng ta dùng từ “convolt” :V

Diagram

Description automatically generated

một ví dụ mà ranh giới giữa 2 màu đc làm nổi bật lên sau khi “convol” bức ảnh

hay còn gọi là “filter” trong những bức ảnh

vấn đề của convolt là các pixel ở viền ít quan trọng, “đóng góp” sau khi qua filter, và kích cỡ ảnh khi qua filter bị giảm

nên một giải pháp có thể thêm vào là padding với 1 cái viền “cho có” để tránh những thứ ở trên

Diagram, engineering drawing

Description automatically generated

vậy padding thêm vào ra sao cho hợp lí, phía dưới là ví dụ, cũng như cách tính padding sao cho kích kỡ không đổi

Text

Description automatically generated

p là padding, n là kích thước input, từ đó ra cái công thức đóng khung, và 1 note nhỏ là kích cỡ ảnh thường là lẻ

như vậy thì mới chia hết đc còn nếu không thì có thể thêm padding cho thành lẻ =))

một vài convolute khác là thay vì nhảy một ô một giờ có thể nhảy 2 ô một hoặc thậm chí nhiều hơn, cụ thể xem ảnh dưới

A picture containing text, electronics

Description automatically generated

tổng hợp lại công thức tính

Diagram

Description automatically generated with medium confidence

giờ triển khai vào một bức ảnh bình thường, sẽ có 3 lớp R, G, B với mỗi lớp sẽ cần một filter tương ứng

Diagram

Description automatically generated

output sẽ là ảnh đen trắng, 1 lớp = tổng của phép nhân ma trận 3\*3\*3 vs 3\*3\*3, dồn lại về 1 số

mình có thể chỉ apply filter cho 1 chanel RED, mà k có GREEN hay BLUE, hoặc có thể apply cả 3

Multiple filters: khi áp nhiều hơn 1 filter thì dồn chúng lại bằng cách stack theo chiều ngang thành nhiều chanel

đây là lí do vì sao mà qua 1 filter thì ảnh về 1 chanel, điều này để tiện cho việc apply nhiều filter

Diagram

Description automatically generated

mình sẽ add convolution layer này vào mạng của mình

Diagram

Description automatically generated

ở ví dụ trên thì qua layer “conv” đầu tiên thì a[1] có kích thước 4\*4\*3

summary

Text, letter

Description automatically generated

simple convolution network

Diagram

Description automatically generated with low confidence

types of layer in convolutional network

- convolution: conv

- pooling: pool

- fully connected: FC

max pooling: từ lấy max trong một hình vuông filter

A picture containing diagram

Description automatically generated

pooling layer này không có parameters nên không thể tối ưu nếu dùng grad decent

và do vậy nếu đặt chung vs các layer khác thì không được tính là một layer của mạng

Diagram, engineering drawing

Description automatically generated

Table

Description automatically generated

một ví dụ về full CNN, với nhiều loại layer trong một mạng, từ conv, đến pool, và full connected: cái FC này chính là mạng neurol đã học từ trước

Why Convolution

Diagram

Description automatically generated with medium confidence

lý do đầu tiên dễ thấy là nếu như dùng mạng neuron đã học từ trước thì những layer đầu sẽ siêu siêu nhiều (hàng tỉ)

parameter vì chúng là full-connected + một bức ảnh tầm 1k pixel là toang hệ thống

và qua đó mình cũng thấy đc convolution có thể kéo số lượng param về mức chấp nhận được và với mọi ảnh độ phân giải mình có thể apply filter tùy chỉnh để kéo dữ liệu về mong muốn