P

7-9-2021 Tuesday

week#2 lecture #2

Image classification

a core task in Computer Vision

<u>The Problem:</u>

Semantic gap

for the computer it is just a tensor of integers b/w [0,255]

e.g $800 \times 600 \times 3$ 3 channels of RGB

2 types of images

a) Gray scale images — Gray shades 0, 255 → Dark pixel, white pixel — single matrix

b) RGB Images, colored images — 3 matrices 2d or 3D matrix

$2^8 \times 2^8 \times 2^8 = 2^{24}$

<u>Challenges:</u>

① Viewpoint variations

All pixels change when the camera moves

② Background Clutter

Background & foreground would have almost same numbers hence Colors

③ Illumination

④ occlusion

Some part of the object is visible only in the image

⑤ Deformation

Daily life objects shape isn't rigid, so shape changes we have to detect it.

⑥ Intraclass variation

An Image classifier:

Unlike e.g a Sorting a list of numbers.

we can't hardcode the algorithm for an object

Attempts have been made:

Find edges

Find corners

ML: Data driven Approach:

① Collect a dataset of images & labels

② Use ML algo's to train a classifier.

③ Evaluate the classifier on new images

Nearest Neighbor classifier:

First classifier:

Nearest Neighbor

2 Steps

① Memorize all data & labels

② Predict the ~~label~~ label of most similar training image

Distance Metric

Example Dataset: CIFAR10:

10 classes

50,000 training Images

10,000 Testing Images

Distance Metric:

To compare Images

L1 distance :

$$d_1(I_1, I_2) = \sum_p |I_1^p - I_2^p|$$

Take Pixel wise difference absolute value & add them

minimum difference = 0
max difference = 255 x dimensions of Pictures

Mostly X means data or images
y means labels
tr = training set

N = no. of images
D = Dimensions
X = N x D
↓
it should be
$N \times D_1 \times D_2 \times 3$
RGB
D

Training = O(1)
Prediction = O(N) for 1 image
O(N^2) for N images

Problem = for Large N images Prediction takes more time.

we want faster Prediction & Slow training is fine for Now.

K-Nearest Neighbors:

we could have noise in data & our Prediction could be wrong so NN has extension K-NN.

Take odd K for voting in NN's and most votes have the label as our Prediction