PA 2018 -01

Recap: PR

结构框图

- Recording(ϕ)
  - Sampling
  - quantization
- Preprocessing(ϕ')
  - smoothing
  - denoisiong
  - Histogram Equlazation
- Fearure Extraction(c)
  dim(f)>>dim(c)
  - edge filter
  - HOG
  - SIFT/SURF
- Classification
- Training

Interlude: Deep Learining

DL needs more data.
eg. 1/10000 hard to learn

low generise

overfitting

underfittting

100*100 data set class dog

75% train data, 25% test set

- count determine relative frequences of correct predictions
  eg accucy recall F1-Score

ROC croves

data $[0, 1]$ label = $[dog, notdog]$
ROC偏向左上的效果最好，对角线的是随机猜测。

Use AUC(area under the curve) to evalue it

Three point

Strict separation of test/train data

- foldas train/test

Cross validation

do not use train data to test

Overfitting

train to well