

Image Classification

1. Semantic Gap : image를 computer가 이해하는 방식은 그저 RGB pixel 값.

⇒ img를 컴퓨터가 이해하는 것과 실제 img 사이에 Semantic gap이 존재

2. Challenge

- ① Background clutter : bg color와 사물의 color가 비슷
- ② Illumination : 빛의 강도, 양에 따라 다르게 보임
- ③ Occlusion : object 일부가 가려진
- ④ Deformation : 일반적인 형태에서 벗어난
- ⑤ Interclass variation : e.g. 고양이 내에서도 종이 여러가지 존재 ⇒ 여기서 오는 특성의 차이까지 감안하는 모델이 필요

3. ML Approach

- ① Collect a dataset of images and labels : dataset 준비
- ② Use ML Algorithm to train a classifier : model 학습
- ③ Evaluate the classifier on new images : model 성능 평가

4. Nearest Neighbor Classifier

- ① train : memorize all data and labels $O(N)$
- ② test : predict the label of the most similar training image $O(N)$

← 우리가 원하는 것: train이 오래 걸리더라도 test가 빠른
이건 완전 반대

- Distance Metric

| Query img, train img | → R

(1) L1 distance : $d_1(I_1, I_2) = \sum_p |I_1^p - I_2^p|$ → L2에 비해 수직/수평 경계가 만들어지는 경향

(2) L2 distance : $d_2(I_1, I_2) = \sqrt{\sum_p (I_1^p - I_2^p)^2}$

③ K-Nearest Neighbor : k는 hyperparameter → Best value of k?

- problem-dependent
- 다해보아야 함.

hyperparameter : 사용자가 직접 값을 지정하는 parameter

ex) k-Nearest Neighbor의 k,

distance metric

train, val, test set으로 나누서 validation에서 최적의 K 값을 찾는 것.

→ dataset이 작다면? Cross-validation

fold 1	fold 2	fold 3	fold 4	fold 5	test
fold 1	fold 2	fold 3	fold 4	fold 5	test
fold 1	fold 2	fold 3	fold 4	fold 5	test
fold 1	fold 2	fold 3	fold 4	fold 5	test
fold 1	fold 2	fold 3	fold 4	fold 5	test

train
val
test

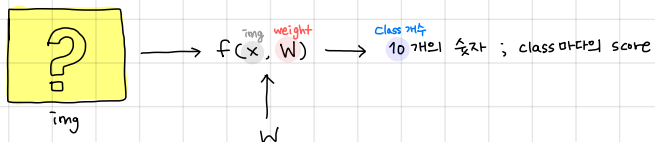
→ 이렇게 한 다음에
결과를 avg 함.

→ small dataset에는 좋지않지만,
DL에서는 잘 사용되지 않음.

④ 단점

- (1) pixel 단위로 얻어낸 distance가 유의미하지 않음.
- (2) test time이 매우 느림
 - ⊕ dim이 늘어남에 따라 point 개수는 기하급수적으로 늘어남.

5. Linear Classifier - Parametric Approach



$\Rightarrow f(x, W) = Wx + b$

 $W: 10 \times 3072$ (weight)

 $x: 3072 \times 1$ (input vector)

 $b: 10 \times 1$ (bias)

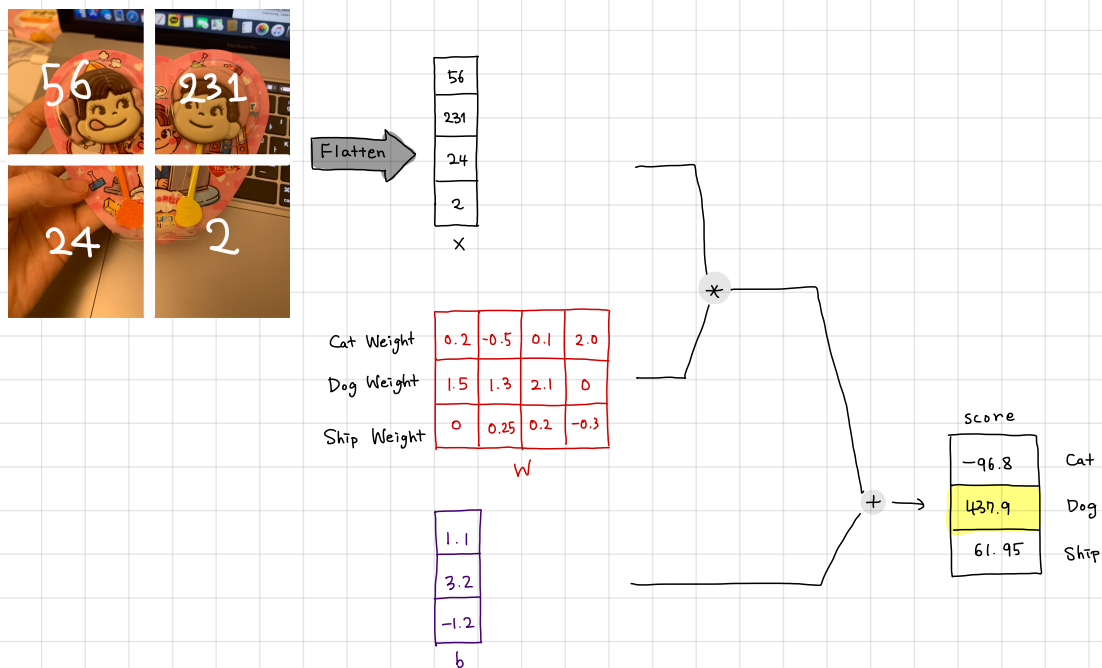
 $Wx: 10 \times 1$

 $Wx + b: 10 \times 1$

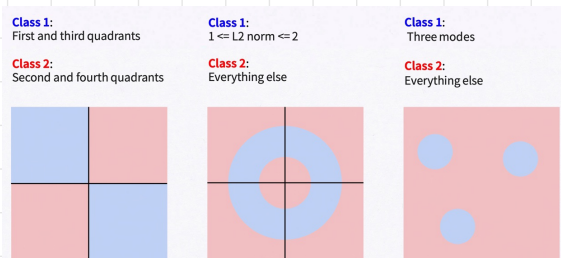
 \Rightarrow 이렇게 나온 score를 통해 label을 예측하고, 이 예측이 잘 맞도록 W, b 를 조정하는 것!

이런 linear classifier를 여러 개 쌓아서 Neural Network를 만들 수 있다.

ex)



① 단점 : 비선형성이 없어 다음과 같은 경우 classify하기 힘들



② Viewpoints

