Lect 12 Face Detection

Face detection & Recognition.

↙ only find the face ↘ know who it is

a mature technology

Challenge for face detection.

- sliding window should cover all possible locations and say "True" or "False"

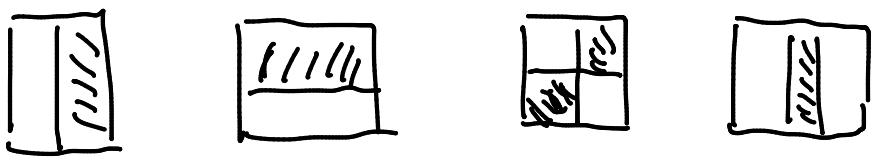
But faces are rare. Say. in a image with 1Million pixels, only at most 10 faces are common.

- Time efficiency. Spend less time on non-face areas
- False positive / False alarm rate should $< 10^{-6}$

Viola/Jones Face Detection.

- use integral image for fast feature evaluation
- use boosting method for feature selection
- additional cascade strategy for fast rejection

Feature detector.



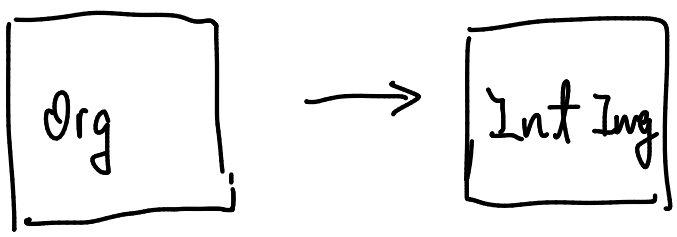
Weak Filters
with essemble method

▨ = -1 ▯ = 1

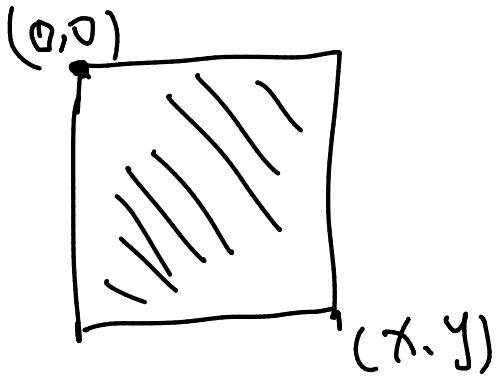
So the value = $\sum$ white area − $\sum$ black area

If image is totally noise, the summation will ≈ 0

We choose these "rectangle" like detectors. because we could make use of integral image

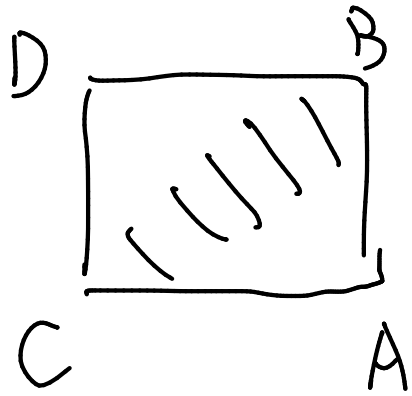


pixel in integral image say (x,y).
is the summation from (0,0) to (x,y)



$$Val(x,y) = \sum_{x,y} img(x,y)$$

the operation could be paralleled. First sum row, then sum col.



$$sum = A - B - C + D$$

So, we could have a efficient filter computor But we still could not deploy thousands of filters of different sizes.

⇓

Boosting / Ensemble classifier.

Weak learners ⇒ Strong ensemble learner.

→ for feature filter i, it's response is h_i

$$h_i(x) = \begin{cases} 1 & \text{if } p_i f_i(x) > p_i \theta \\ 0 & \text{else} \end{cases}$$

$h_i(x)$: window; $f_i(x)$: filter result; p_i: parity; θ: threshold

→ Ensemble Classification function

$$C(x) = \begin{cases} 1 & \text{if } \sum_{i=1}^{T} \alpha_i h_i(x) > \frac{1}{2} \sum_{i=1}^{T} \alpha_i \\ 0 & \text{else} \end{cases}$$

half vote; majority vote; learnt weight

Training Stage.

- Initially, each training example are equally treated
- For each boosting round:
  - → Find the weak learner achieves lowest weighted training error
  - → Raise the misclassified training examples' importance

  Intuation. we find out which weak learner that can do well on difficult examples
- Final classifier is a combination of all weighted weak learners

---

Boosting:

Pros:

① Testing is fast ② Easy to implement ③ flexible of weak learner choosing

Con:

① Traing slow ② needs lots of samples

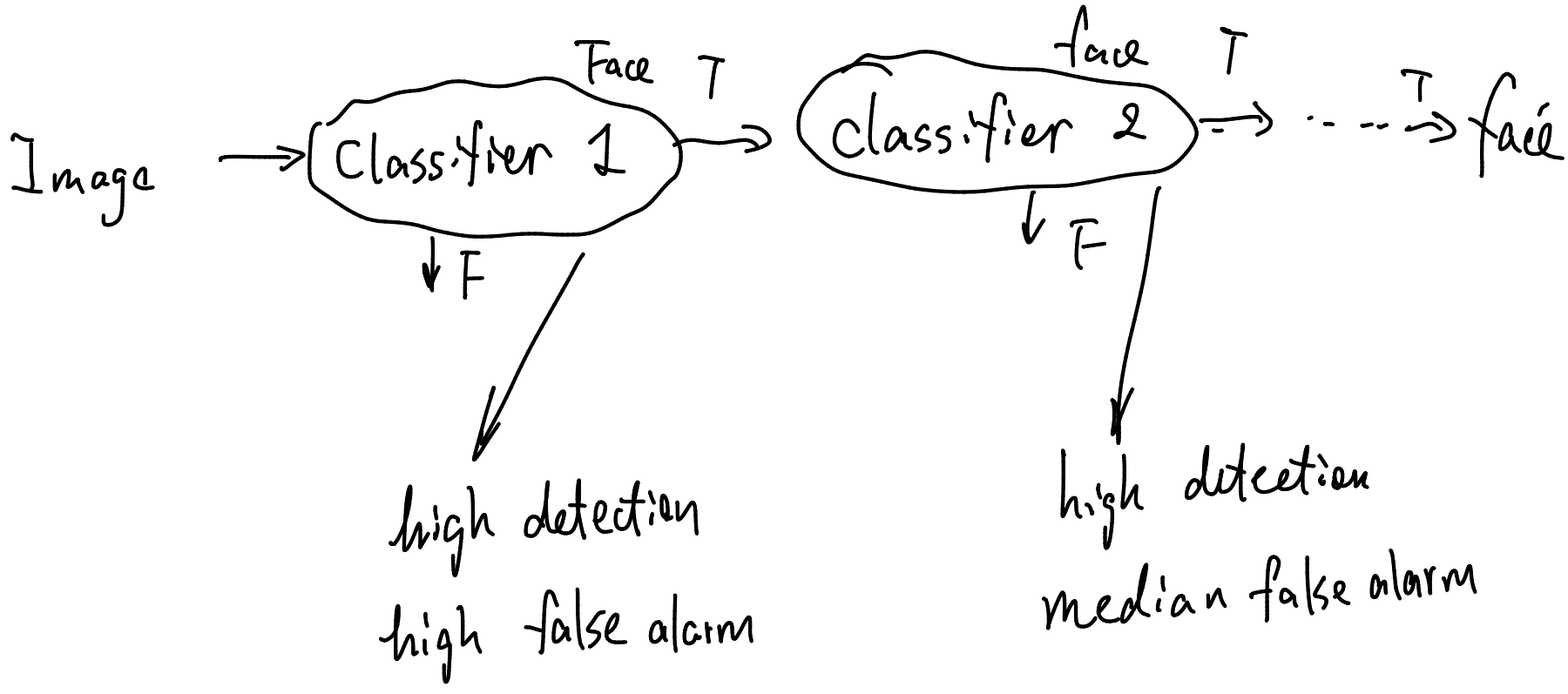
ROC (Receiver Operating characteristic) Curve.

a 200-feature detecact would have 95% Detection / $1/14_{089}$

False alarm.

The false alarm rate is high. 'ecause we may apple the filter millions of time. We need more!

We have another fiter could reject non-faces.



We can apply more and more complecated filters.

Train cascade

- Set a target detection/false alarm rate
- keep adding features, until the target rate is met
- Add stage if false alarm rate is not met

Training set

say 1% face, 99% non-face