Week 11

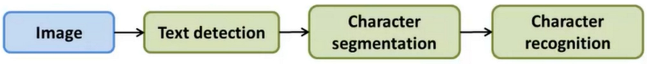
- Photo OCR:
  a. Problem Description and Pipeline:

  Photo OCR: Photo Optical Character Recognition

  First, given a picture it has to look through the image and detect where there is text in the picture.

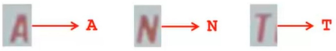
Photo OCR pipeline

1. Text detection



2. Character segmentation

3. Character classification

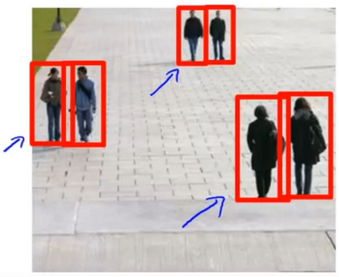


  b. Sliding Windows:

  these rectangles that you're trying to find can have different aspect

Text detection

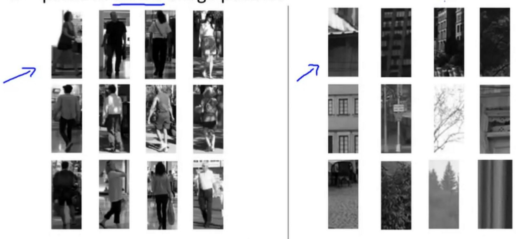
Pedestrian detection



Supervised learning for pedestrian detection

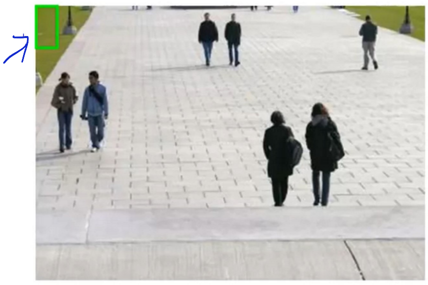
$x =$ pixels in 82x36 image patches

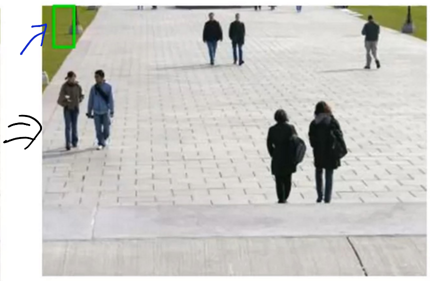
1000
10,000

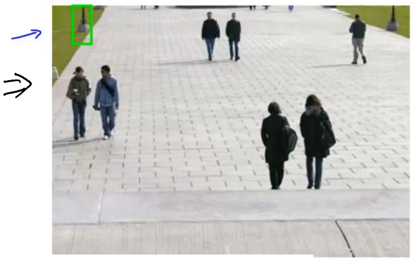


Positive examples $(y = 1)$ Negative examples $(y = 0)$

step-size / stride



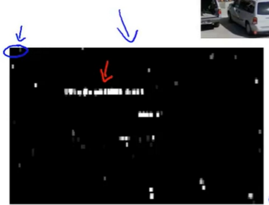




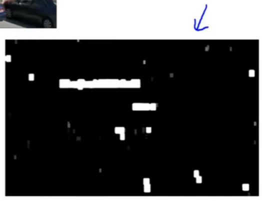
82x36 take a larger patch and resize to 82x36.

Text detection

"expansion"

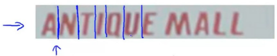


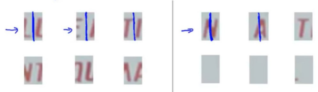
expansion operator.



grow the white pixels a little by seeing whether the nearby pixels.

1D Sliding window for character segmentation





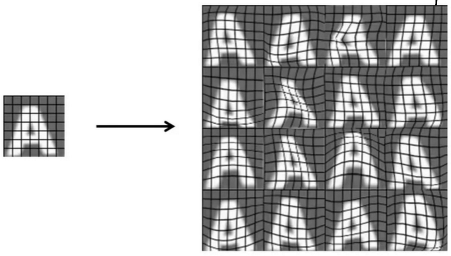
Positive examples $(y = 1)$ Negative examples $(y = 0)$

c. Getting Lots of Data and Artificial Data:

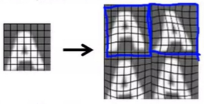
Artificial data synthesis

create from begining.
already have a small label training set and we amplify training set.

Synthesizing data by introducing distortions

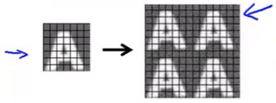


→ Distortion introduced should be representation of the type of noise/distortions in the test set.



→ Audio:
Background noise,
bad cellphone connection

→ Usually does not help to add purely random/meaningless noise to your data.



→ $x_i =$ intensity (brightness) of pixel i

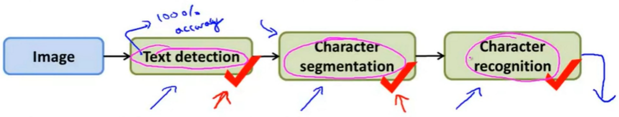
→ $x_i \leftarrow x_i +$ random noise

Discussion on getting more data

1. Make sure you have a low bias classifier before expending the effort. (Plot learning curves). E.g. keep increasing the number of features/number of hidden units in neural network until you have a low bias classifier.
2. "How much work would it be to get 10x as much data as we currently have?"
   - Artificial data synthesis
   - Collect/label it yourself
   - "Crowd source" (E.g. Amazon Mechanical Turk)

d. Ceiling Analysis What Part of the Pipeline to Work on Next:

Tell you what parts of the pipeline might be the best use of your time to work on.



What part of the pipeline should you spend the most time trying to improve?

Component	Accuracy	overall system.
Overall system	72%	
→ Text detection	89%	↓17%
Character segmentation	90%	↓1%
Character recognition	100%	↓10%

how much could you possibly gain if one of these components became absolutely perfect?