

# WEEK - 11

## ⇒ Photo OCR Pipeline

- 1) = Text detection
- 2) = Character Segmentation
- 3) = Character Classification or Recognition

## ⇒ Sliding Window Detection :-

- We take an image and extract patches of it. A patch would be of a fixed size for iterating an image once. Then the patch size ~~is~~ is increased and the patch is iterated over the image.
- By this method, we can detect different pedestrians <sup>at</sup> of different distances from the camera which makes them of diff sizes.
- The patch ~~is~~ / window is shifted by a small amount of pixels (4 or 6 pixels) called the step size or Stride.
- Every patch of whatever size is resized to a certain ~~length~~ height & width ratio that remains constant throughout.



⇒ Expansion operator → It takes the white regions and it expands them.

⇒ Artificial Data Synthesis for OCR

1) = Taking online fonts and pasting them in a new background

2) = Synthesizing already acquired data by introducing distortions to increase the dataset.

⇒ Make sure you have a low bias classifier before expanding the effort. Eg. keep increasing the number of features / number of hidden units in neural network until you have a low bias classifier.

⇒ Ways of collecting data :-

1) = Artificial Data Synthesis

2) = Collect / label it yourself

3) = Crowd Sourcing (Amazon Mechanical Turk)

⇒ Cacloring Analysis :- We go through every stage ~~step~~ of our pipeline and give each stage 100% accuracy. And after that we check the overall accuracy of the system. The difference in accuracy of the stage of the pipeline ~~gives~~ tell us its importance.

- ⇒ We see what affect each stage causes on the final accuracy ie. we determine their importance.
- ⇒ And based on their importance, we allocate our time & resources to that stage.
- ⇒ The stages that contribute largely towards increasing the overall accuracy become important and ~~no~~ vice versa.

