

photo OCR (photo optical character recognition).

how to let computers read text to the pixel in images.

① Text detection: find the region where the text is.

② Character segmentation: segment individual character.

③ character classification: recognize each character

Sliding windows.

pedestrians detection

$\left\{ \begin{array}{l} \text{positive examples (y=1)} \rightarrow \text{pedestrians} \\ \text{negative examples (y=0)} \rightarrow \text{others.} \end{array} \right.$

Using a window with different aspect of ratio and sliding it over the whole image to detect the objects (pedestrians), sending them to the classifier.

Sliding window for character segmentation:

$\left\{ \begin{array}{l} \text{positive examples (y=1)}: \text{the middle of image represents a gap or a split} \\ \text{negative examples (y=0)}: \text{single character, no need to split.} \end{array} \right.$

Getting lots of Data and Artificial Data. (low bias classifier)

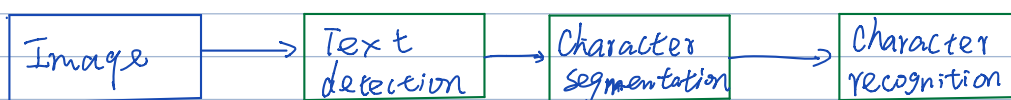
① Create a new dataset from the beginning based on specific application.

② Using raw training sets, modifying them into new training examples, such as distortion, blurring, rotation. (data augmentation)

the distortion introduced should be representation of the type of noise/distortion in the test set. (Should be reasonable).

Ceiling Analysis:

Using ceiling Analysis to decide which part we should spend more time on.



1. 如果系统本身的正确率为72%.

Component	Accuracy
Overall system	72%
Text detection	89%
Character segmentation	90%
Character recognition	100%

2. 如果令 Text detection 的输出 100% 正确, 发现系统的总体效果从 72% 提升到了 89%, 这意味着我们可以花点时间来提高 text detection 的精度。

3. 如果我们手动选择数据来进行 Character segmentation, 使其输出 100% 正确, 但系统总体效果只提升了 1%, 这说明 Character segmentation 已经足够好了。

4. 在 Character recognition 阶段采用同样的方法, 发现系统效果提升了 10%, 那我们不需要花时间来提高名用的总体效果。