16.吴恩达-机器学习+应用举例-照片OCR

笔记本: 日常
创建时间: 2019/11/18 9:13 **更新时间:** 2019/11/18 11:10
作者: 296645429@qq.com

Application example: Photo OCR

Problem description and pipeline

Photo OCR pipeline

→ 1. Text detection

→ 2. Character segmentation

→ 3. Character classification

Andrew

滑动窗口

Application example: Photo OCR

Sliding windows

Supervised learning for pedestrian detection

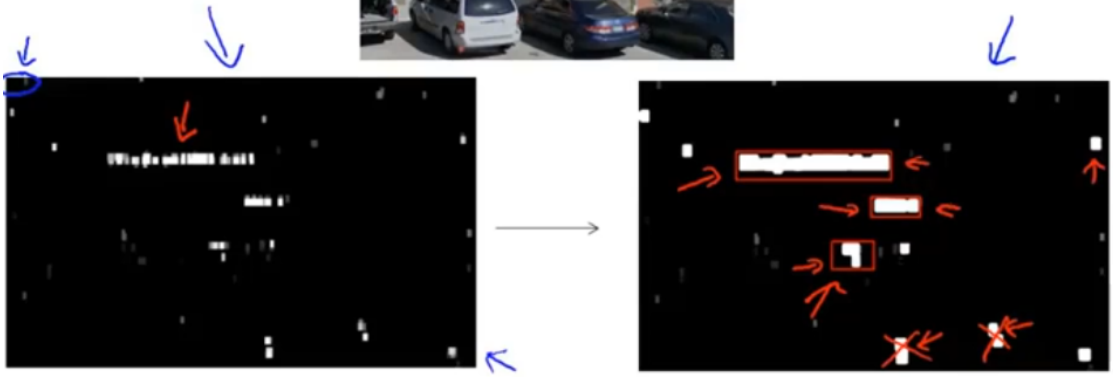
$x =$ pixels in 82x36 image patches

Positive examples $(y = 1)$

Negative examples $(y = 0)$

Text detection

"expansion"



1D Sliding window for character segmentation

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

Photo OCR pipeline

1. Text detection

2. Character segmentation

3. Character classification

获取数据的方法

Application example: Photo OCR

Getting lots of data: Artificial data synthesis

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)

Application example: Photo OCR

Ceiling analysis: What part of the pipeline to work on next

上限分析

Estimating the errors due to each component (ceiling analysis)

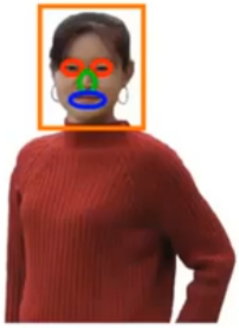
What part of the pipeline should you spend the most time trying to improve?

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

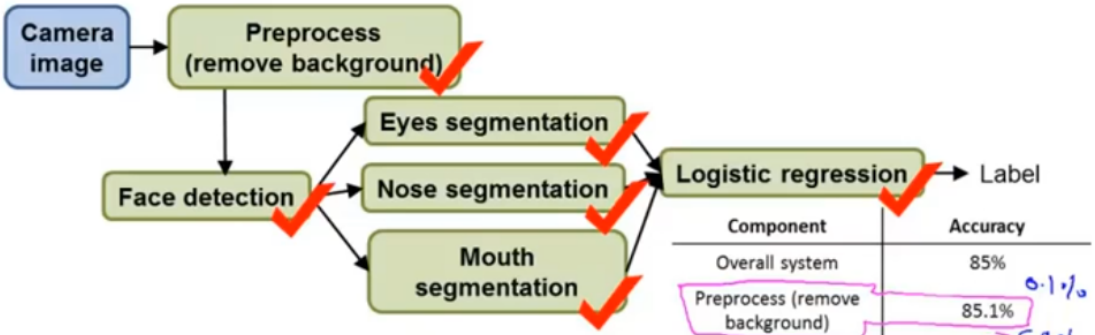
↓17%
↓1%
↓10%

Another ceiling analysis example

Face recognition from images
(Artificial example)



Another ceiling analysis example



Component	Accuracy
Overall system	85%
Preprocess (remove background)	85.1%
Face detection	91%
Eyes segmentation	95%
Nose segmentation	96%
Mouth segmentation	97%
Logistic regression	100%

0.1%
5.9%
4%
1%
1%
3%

Andrew Ng