Photo OCR pipeline

1. Text detection

2. Character segmentation

3. Character classification

Photo OCR pipeline

1. Text detection

2. Character segmentation

3. Character classification

C1eaning → Cleaning

Photo OCR pipeline

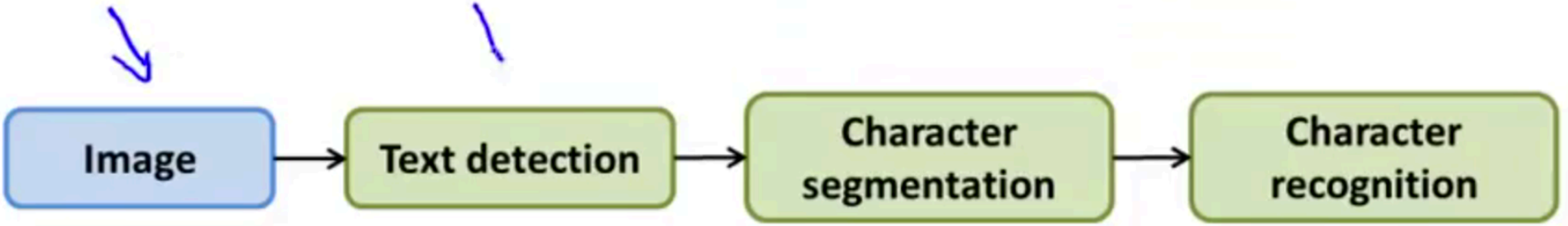
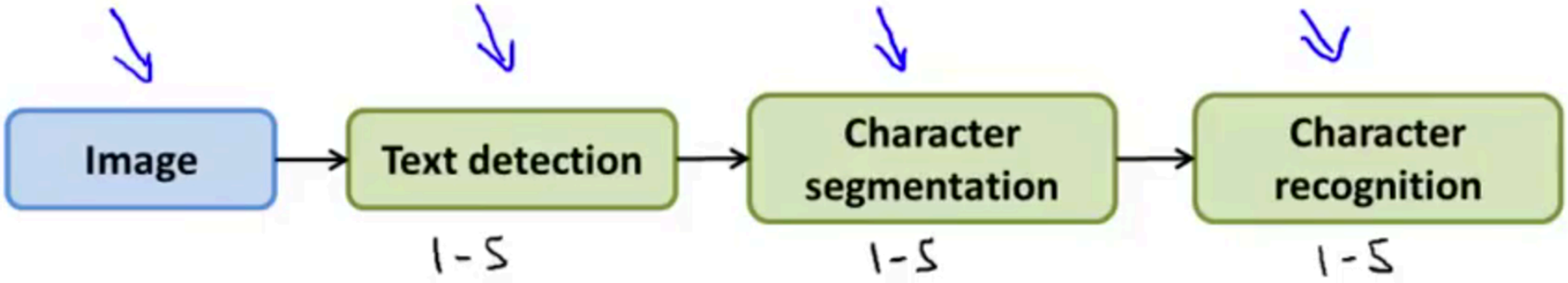
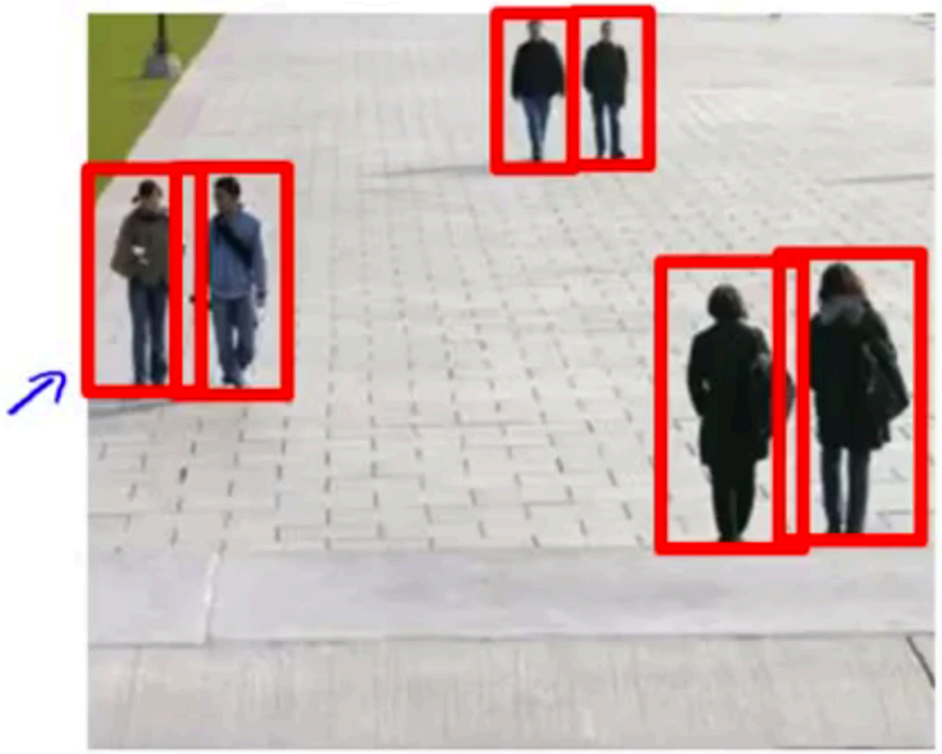


Photo OCR pipeline



Text detection

Pedestrian detection



Andrew Ng

Supervised learning for pedestrian detection

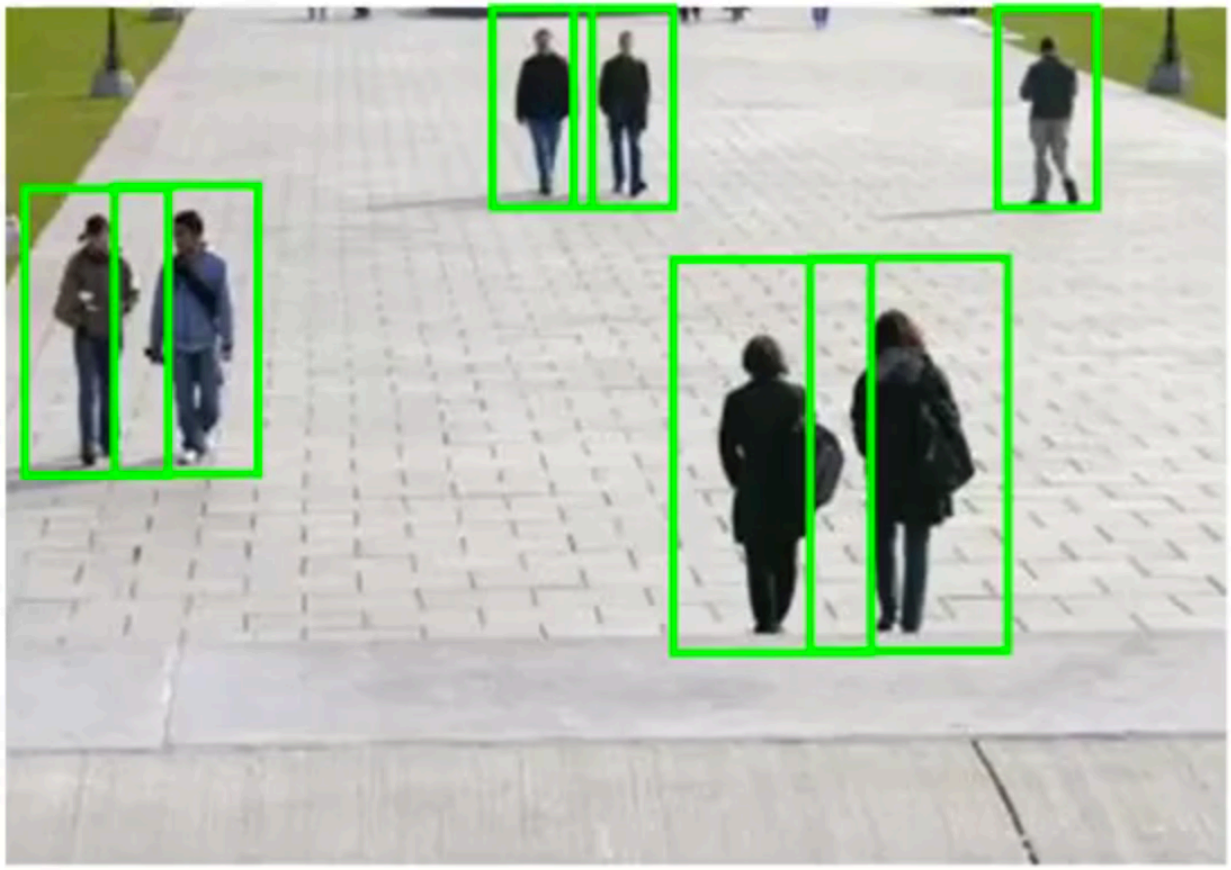
$x =$ pixels in 82x36 image patches

Positive examples $(y = 1)$

Negative examples $(y = 0)$

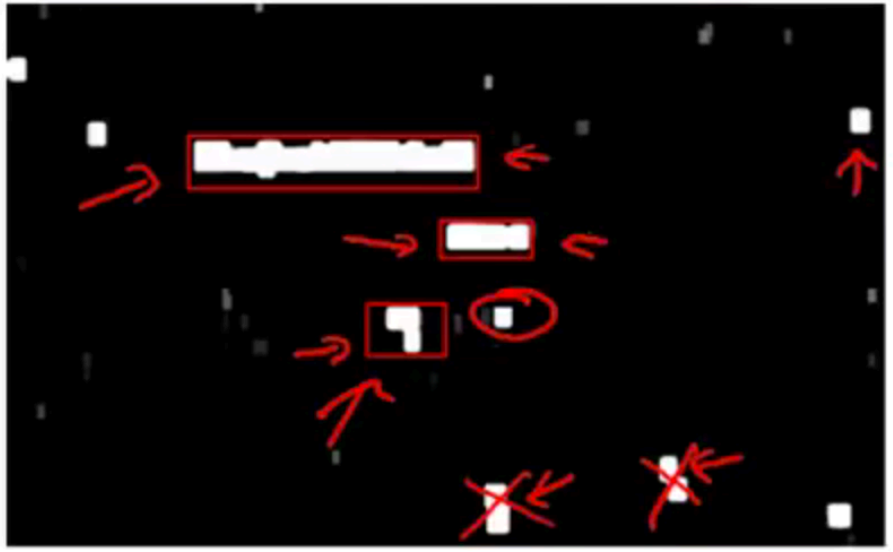
Andrew Ng

Sliding window detection



Text detection

"expansion"

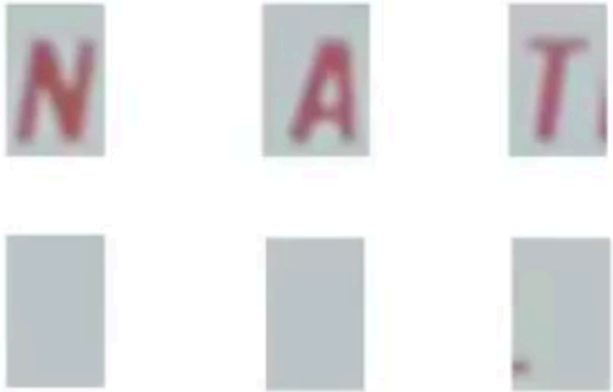


[David Wu]

Andrew Ng

1D Sliding window for character segmentation

Positive examples $(y = 1)$



Negative examples $(y = 0)$

Artificial data synthesis for photo OCR

Real data

[Adam Coates and Tao Wang]

Artificial data synthesis for photo OCR

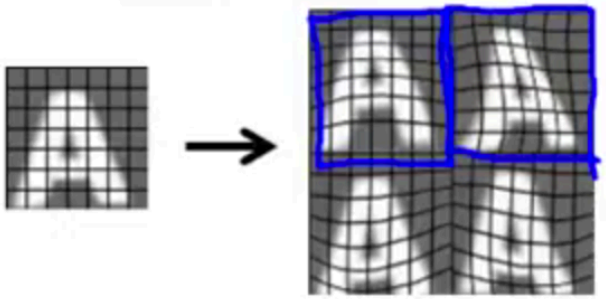
Real data

Synthetic data

[Adam Coates and Tao Wang]

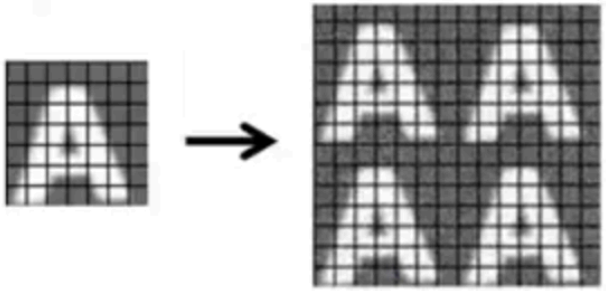
Synthesizing data by introducing distortions

- 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

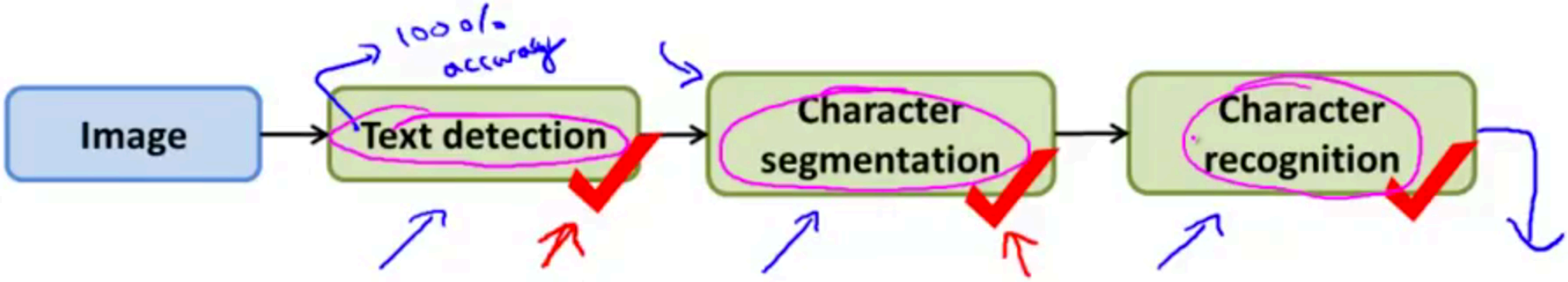
$x_i \leftarrow x_i +$ random noise

[Adam Coates and Tao Wang]

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)

Estimating the errors due to each component (ceiling analysis)



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

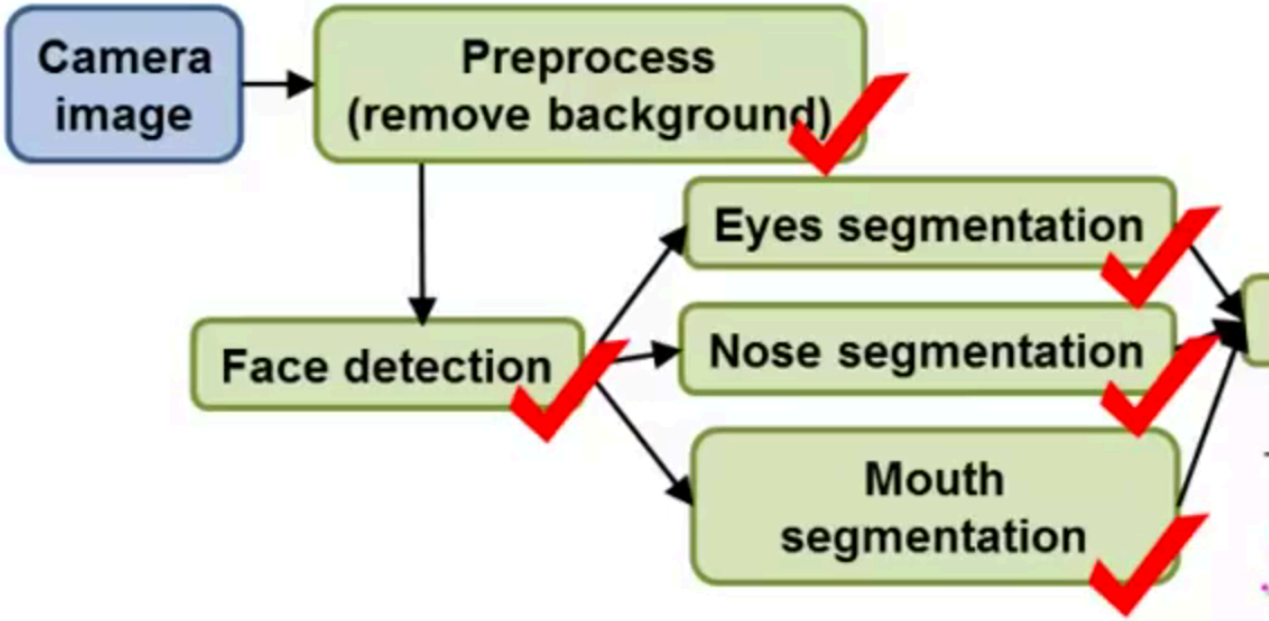
Component	Accuracy
Overall system	72%
Text detection	89%
Character segmentation	90%
Character recognition	100%

↓ 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

Summary: Main topics

Supervised Learning
- Linear regression, logistic regression, neural networks, SVMs

Unsupervised Learning
- K-means, PCA, Anomaly detection

Special applications/special topics
- Recommender systems, large scale machine learning.

Advice on building a machine learning system
- Bias/variance, regularization; deciding what to work on next: evaluation of learning algorithms, learning curves, error analysis, ceiling analysis.

Summary: Main topics

$(x^{(i)}, y^{(i)})$

- Supervised Learning
  - Linear regression, logistic regression, neural networks, SVMs

$x^{(i)}$

- Unsupervised Learning
  - K-means, PCA, Anomaly detection

- Special applications/special topics
  - Recommender systems, large scale machine learning.

- Advice on building a machine learning system
  - Bias/variance, regularization; deciding what to work on next: evaluation of learning algorithms, learning curves, error analysis, ceiling analysis.