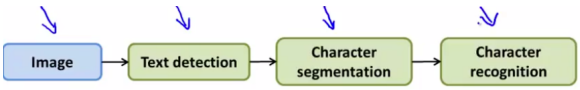
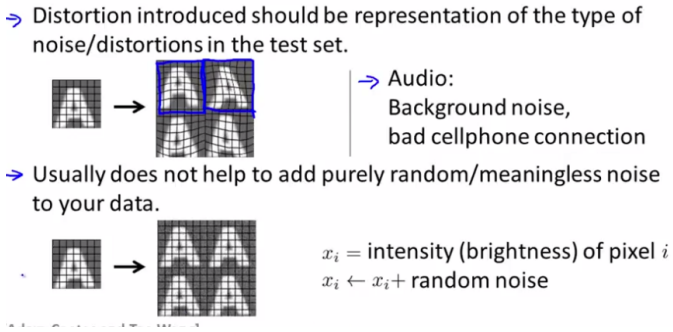
- **Photo OCR**
- **Problem Description and Pipeline**
- Photo OCR pipeline Optical character recognition
  - Text detection
    - find the regions of where the there is text in the image
  - Character segmentation
    - try to segment the text into individual characters
  - Character classification
    - looks at the images of the individual characters and figure what each character represents
- Photo OCR pipeline
  - how to break the system down into different modules

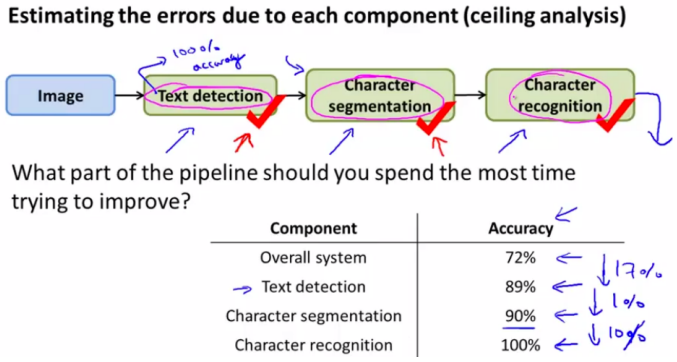


- **Sliding Windows**
- Text detection (detection in anything example pedestrian detection)
- Supervised learning for pedestrian detection
  - x = pixels in 82x36 image patches
  - positive examples (y=1)
  - negative examples (y=0)
- Sliding window detection
  - take a rectangle or window and slide over the image using a step size of stride size
- Text detection
  - positive examples (y=1)
  - negative examples(y=0)
  - white regions show where the classifier thinks it found text
    - take the white regions and perform an expansion to that region
      - is a pixel within a range of a white area in the classifier output
      - to obtain the new white regions that are expanded
- 1d sliding window for character segmentation
  - positive examples (y = 1)
  - negative examples (y = 0)
- **Getting Lots of Data and Artificial Data**
- Character recognition
  - takes as input an image and recognize what character it is
- Artificial data synthesis for photo OCR
  - take characters from different fonts and paste characters against random backgrounds
  - perform scaling and other things to obtain synthetic data
  - using synthetic data unlimited supply of training data
- Synthesizing data by introducing distortions
  - introduce distortions to obtain more training data
  - ways to amplify or multiply training set
- Synthesizing data by introducing distortions: Speech recognition
  - add background noise to the training examples to gain more training data
- Synthesizing data by introducing distortions
  - Distortion introduced should be representation of the type of noise/distortions in the test set

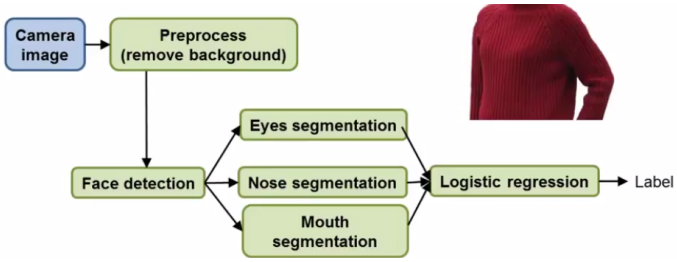
  - Audio: Background noise, bad cellphone connection
- Usually does not help to add purely random/meaningless noise to your data



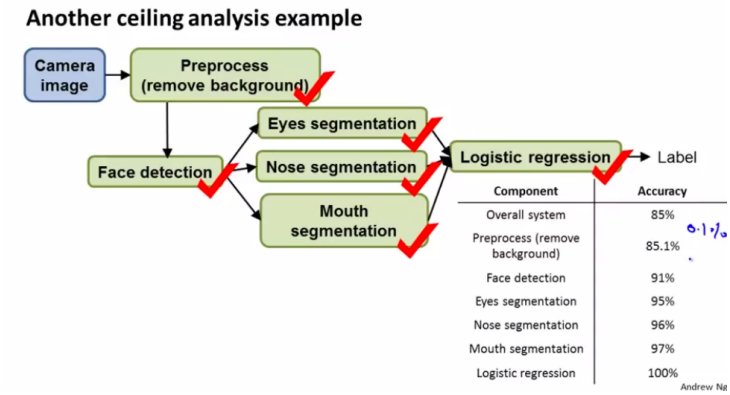
- Discussion on getting more data
  - make sure you have a low bias classifier before expending the effort. (Plot learning curves). E.g. keep increasing the number of features/number of hidden units in neural network until you have a low bias classifier
  - How much work would it be to get 10x as much data as we currently have?
    - artificial data synthesis
    - collect/label it yourself
    - "crowd source" (e.g. amazon mechanical turk)
- **Ceiling Analysis: What Part of the Pipeline to Work on Next**
  - Estimating the errors due to each component (ceiling analysis)
  - image -> text detection -> character segmentation -> character recognition
  - what part of the pipeline should you spend the most time trying to improve?



- Another ceiling analysis example
  - Face recognition from images(Artificial example)



- Another ceiling analysis example



-