

Week 18

August 20, 2021

Problem description and pipeline

- Consider how a complex system may be built.
- The idea of a machine learning pipeline.
- Machine learning applied to real-world issues and artificial data synthesis.

What is the photo OCR problem?

- Photo OCR = photo optical character recognition.
- Getting computers to interpret digital images is one notion that has piqued the curiosity of many people.
- The idea behind picture OCR is to teach computers to interpret text in images.

Pipeline - a sequence of separate modules, each of which may be a machine learning or data processing component.



Pedestrian detection

- We're looking for pedestrians in the picture.
- A common aspect ratio for a standing human is 82 x 36.
- Obtain a training set of both positive and negative examples (1000 - 10 000).



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

- Now that we have a new image, how do we identify pedestrians within it?
- Start by taking a rectangular 82 x 36 patch in the image.
- Run patch through classifier (returns 1 or 0).
- After that, move the rectangle slightly to the right and re-run the program.
- Repeat until you've covered the whole image.
- Hopefully, by varying the patch size and rastering over the image frequently, you will finally recognize all of the pedestrians in the image.

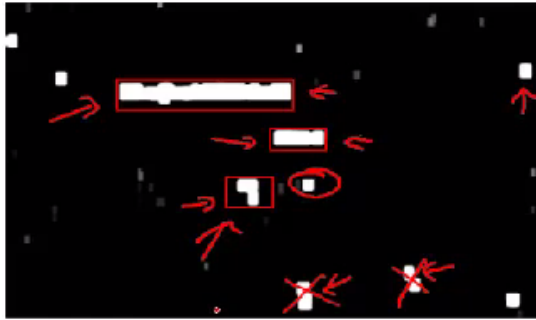


Text detection example

- We generate a labeled training set with positive examples (some type of text) and negative examples (not text), similar to pedestrian detection.
- After training the classifier, we apply it to an image using a sliding window classifier with a set rectangle size.
- Obtain a training set of both positive and negative examples (1000 - 10 000).
- The white region indicates where the text recognition algorithm believes text exists, while the varying shades of gray correlate to the likelihood associated with how certain the classifier is that the section includes text

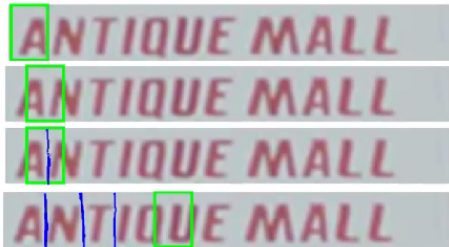


- Take the classifier output and apply an expansion algorithm that extends each of the white areas.
- Examine the linked white patches in the picture above. Draw rectangles around those that make sense as text (tall narrow boxes don't).
- This example misses a piece of text on the door because the aspect ratio is wrong.



Stage two is character segmentation

- To navigate along text areas, use a 1-dimensional sliding window.
- Does each window snapshot resemble the separation of two characters?
- Insert a split if yes.
- If not, proceed.



Character recognition as an example of data synthesis

- The goal is for the system to recognize a character based on an image patch.
- Consider the photos to be grayscale (makes it a bit easier).

- Where can I find training data?
- Modern computers frequently include a large font library, and there are several free font libraries available on the internet.
- Take characters from different fonts and place them on random backgrounds to get more training data.
- Another approach is to add distortion into existing data, such as warping a character.



Synthetic data