Predicting:

This project concentrates on recognizing medicine name from their tablet. at first, we plan on applying the YOLOV3 detector for detecting the medicine name outline, then use the result and feed it into the CRNN for text recognition

Yolo detector

CRNN

text

After apply yolo

image

Data:

our data comes from two different resources: The first includes photos collected from pharmacy, the second includes photos collected from the web scraping

in aggregation, we train YOLOV3 detector and CRNN on 4000 real unique photos in different resolutions and variety of extents

\*sample of data

Model:

YOLO, which means that an image can be predicted what the object is and where they are at one glance. it frames object detection as a regression problem, a single convolutional network simultaneously predicts multiple bounding boxes and class probabilities for those boxes, it divides the input image into an S\*S grid and runs the classification and localization problem to each of the S\*S grid simultaneously then use non-max suppression to generate final prediction.

in this project we use YOLOv3, the latest variant of a popular object detection algorithm YOLO. The published model recognizes 80 different objects in images and videos, Moreover, you can easily tradeoff between speed and accuracy simply by changing the size of the model, no retraining required!

* **Bounding Box Predictions** in YOLOv3 gives the score to the objects for each bounding box. It uses logistic regression to predict the objectiveness score.
* **Class Predictions**in YOLOv3 it uses logistic classifiers for every class instead of SoftMax which has been used in the previous YOLOv2
* YOLOv3 makes predictions similar to the FPN where 3 predictions are made for every location the input image and features are extracted from each prediction
* the predecessorYOLOv2 used Darknet-19as feature extractor and YOLOv3 uses the Darknet-53 network for feature extractor which has 53 convolutional layers

