CSC420 Project Proposal

Nutrition Label Extractor

David Shan & Wilbert Guo

Background:
Given a picture of food packaging with a nutrition label, we will extract the information from the nutrition label and output the information in a pie chart showing the breakdown of the macro-nutritional information.

Challenges:

- Not all food packaging is flat, they may be in a cylindrical shape etc.
- Some food packaging color could create difficulties in extracting the label info clearly
- Bad lighting and illumination in photos could cause difficulties in extracting the label

Outline:

Step/ Motivation	Implementation
Image Processing & Label Extraction: We want the image in a clean format where the English text and numerical values are easily readable.	• Convert image to grayscale and apply a blur • Use TSAW skew detection to deal with irregular packaging • Apply histogram equalization to darken nutrition label border • Use Gaussian thresholding to make dark pixels to black and lighter pixels to white • Use contour identification to locate where the nutrition label is located in the image • Apply perspective transform to orientate label and output new image containing only the nutrition label
Information Extraction: We want to take the image generated in the previous step and identify the English words and numerical values	• Use Optical Character Recognition (OCR) • Filter through the Tesseract output and decide which information to keep • Make a key/value pairs data structure with the nutrient name and value • Use bipartite graph matching algorithm to determine the Levenshtein distance between each key to a known nutrient name and minimize total edit distance among all matches • Output information to a text file for making pie charts in the next step
Displaying Information in a Pie Chart: We want to take the information we have extracted and show it in a pie chart for the user to visualize	• Extract the information from the text file produced in the previous step and convert it to python data structures • Use Matplotlib to plot data into pie chart for users to visualize

Citations:

https://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=8204&context=etd
https://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=5916&context=etd
https://people.cs.pitt.edu/~kirk/cs1501/Pruhs/Spring2006/assignments/editdistance/Levenshtein%20Distance.htm