Project Proposal

November 1, 2022

Group Members: Abhaya Shrestha (ashrestha@mines.edu), and Jason Rosenblum (jhrosenblum@mines.edu)

Description
Today, there are often cases where people get ill or die due to products that are unsafe to consume. Our project will detect expiry label on products (particulary food, and medicine), and determine if it is safe to consume which will prevent people from getting ill or dying.

Approach
The approach we plan to take is to use a mixture of techniques from class to achieve the most accuracte result:

1. **Cleanup step:** first cleanup any distortions using camera calibration techiniques (example: the label could be a rectangle in a cylindrical product, or could be on the side of a jar cap; therefore, this step is absolutely crucial) to map the label onto a rectangular image.
2. **Detection step:** For each unique colors of data we collect (see **Data** section) that labels are written in do the following.
    - Use hue, saturation and vue thresholding technique to detect that color, and compute connected components.
    - Convert the connected components of that color to black, with a white background using thresholding.
    - Use the compiled list of keyboard letter templates written in black to match the letters in the processed image (see **Data** section).

- See if the template matching spells any of the starting substring collected in data (see **Data** section). If it does, then that means we have found the substring, and we can use it's geometry to find matched templates to the right of that substring which indicates date.

3. **Evaulation of step:** If the date label could be detected and date format be understood, output whether or not it is safe to consume. Otherwise, if the steps above for detection could not detect it or if date format could not be understood, output not detected.

Data

For our algorithm to work properly, we need to collect data on

1. The unique color labels for food and medicine. This will be used in our detection step for HSV thresholding technique.

2. Limiting the templates to the English keyboard letters, we need to gather the letters written in black for the most common fonts that most food/medicine labels are written in.

3. The unique labels substring for expirey labels (example: "used by", "expires by" etc.)

4. A dataset of images of food and medicine labels which contain the label or which do not contain the label (we also need to test true positive, and true negatives). We can gather these from retail stores, or from online pictures that are free to use.

Timeline

- **November 21st:** finish collecting data, finish research on possible inputs our algorithm (mostly going to be the variety of labels we compile and the unique colors of labels our algorithm needs to detect) needs to process, and start working on implementation.

- **November 28th:** have presentation finished, and continue working on implementation.

- **December 9th:** work on optimizing algorithm accuracy.

- **December 12th:** have code finished, and have paper finished.