

Proposal:

Use Computer Vision (CV) and Machine Learning (ML) in a mobile app for fast and accurate recognition of ingredient lists for detection of allergens in packaged foods. Extend the system in phase 2 for estimation of chemical toxicity in skin-care and other labelled manufactured products.

Background:

Around 15 million people in the United States alone are afflicted with some type of food-allergy, and an allergic reaction hospitalizes someone every three minutes. I have firsthand experience with allergic reactions, having food-allergies myself and witnessing family members react to common foods such as wheat, milk, eggs, and nuts. Reading product-ingredient lists in grocery stores is mandatory in order to prevent ingestion of allergens which could put the victim in a life-threatening situation. However, scanning food labels is a tedious process prone to error in an area where there is no margin for error.

Identifying ingredients printed on various types of packaging is a challenging problem given the variation in shapes of those packages (cans, pouches, bottles, boxes, etc), viewing angles, lighting conditions, and font variability. While traditional Optical Character Recognition (OCR) may help recognize text, its reliability under the various conditions listed earlier requires further study. Use of more advanced algorithms involving ML are likely to produce more accurate results but involve considerable complexity.

Solution

I propose to build a mobile app capable of identifying ingredients from images of the general area of a package where such a list is printed. The app alerts a user to the presence of common allergens in packaged-food items. Users can also specify other items they may wish to avoid.

This app can be enhanced in phase 2 to communicate with a database that provides toxicity information of various chemicals found on items such as bottles of shampoo, body lotion and other skin-care products, along with over-the-counter drugs.

The app utilizes a smartphone's camera and processes the image either locally or transmits the image for cloud-based processing when additional compute resources are required for identifying individual ingredients using ML.

Benefits

The proposed app will significantly speed up scanning of long ingredient lists for potential allergens. Often times ingredients have names that may not appear to be allergens; for e.g., "whey" or "casein" can prove to be deadly for those with milk allergies. Humans scanning the list for words like milk or dairy may miss that ingredient as being a milk derivative. The app would therefore reduce errors associated with shopping for those with food restrictions and ensure the safety of those with allergies.

The proposed phase 2 extension alerts users about the presence of ingredients known to be toxic (carcinogenic, for e.g.) in skin-care and other products. For e.g., use of Triclosan in anti-bacterial soaps has been banned by the FDA due to its potential for hormonal disruption, yet it can still be found in oral products such as Colgate's Total toothpaste! Integrating the app with a database in the cloud that provides updated information on the toxicity of various chemicals and food items will allow for more informed and safer purchases of products.