Leveraging Context to Support Automated Food Recognition in Restaurants

Vinay Bettadapura, Edison Thomaz, Aman Parnami, Gregory Abowd, Irfan Essa
College of Computing, Georgia Institute of Technology, Atlanta, USA

IEEE Winter Conference on Applications of Computer Vision

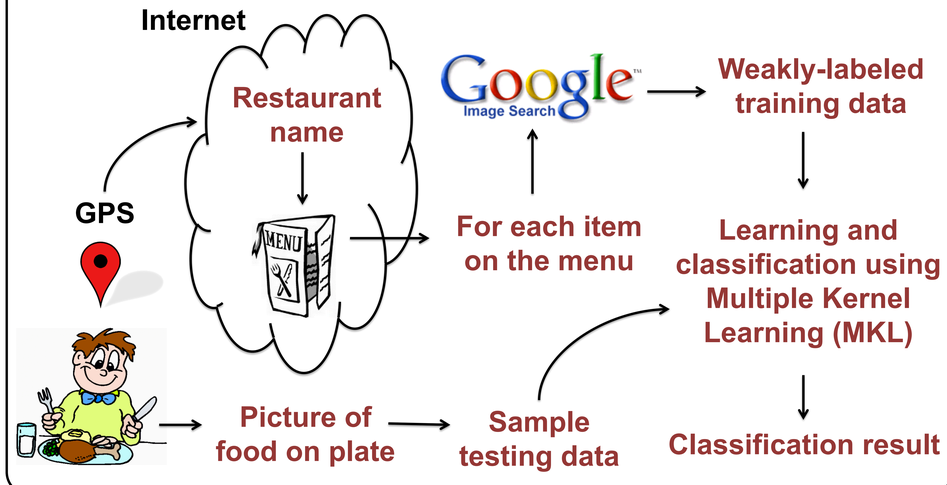
Motivation

- In **1970, 25.9%** of food spending was on food away from home
- By **2012**, this number had risen to **43.1%**
- **80%** of Americans report eating fast-food **monthly** and **40%** report **weekly**
- **Concerns:** Obesity is now a major health concern. Also challenges with nutrition and chronic diseases
- **Solution:** Tracking and logging eating habits using diaries and smartphones
- But manual tracking is **time-consuming**, **prone to errors** and susceptible to **selective under-reporting**

Features & Classifier

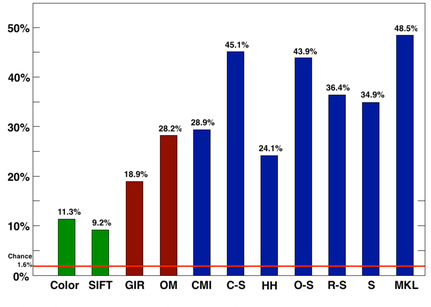
- **Harris-Laplace** point detector followed by **6 feature descriptors:**
  - **CMI:** Color Moment Invariants
  - **C-S:** C-SIFT
  - **HH:** Hue Histograms
  - **O-S:** Opponent SIFT
  - **R-S:** RGB SIFT
  - **S:** SIFT
- **Classifier:** Sequential Minimal Optimization Multiple Kernel Learning (**SMO-MKL**) SVM.

System Overview



Evaluation

- System performance tested on CMU's **PFID** food dataset
- Test images collected across **5 different cuisines** from 10 different restaurants (2 per cuisine): **American, Indian, Italian, Mexican and Thai**
- Total of **600 testing images** (300 using mobile-phone and 300 using Google Glass)
- Total of **3,750 weakly-labeled training images**



Performance of the 6 feature descriptors and SMO-MKL on the PFID dataset.

Evaluation

Weakly-labeled training images obtained from Google Image search for two classes of food. **Left:** Basil Fried Rice. **Right:** Lo Mein.

Sample (12 out of 600) "in-the-wild" images used for testing.

	CMI	C-S	HH	O-S	R-S	S	MKL
American	45.8%	51.7%	43.3%	43.3%	37.5%	29.2%	**67.5%**
Indian	44.2%	74.2%	55.0%	59.2%	69.2%	65.0%	**80.8%**
Italian	33.3%	52.5%	67.5%	**74.2%**	66.7%	49.2%	67.5%
Mexican	36.7%	35.8%	20.8%	37.5%	24.2%	33.3%	**43.3%**
Thai	27.5%	36.7%	25.0%	33.3%	**50.8%**	30.8%	**50.8%**

Classification results showing the performance of the various feature descriptors.

Confusion Matrices