New Technologies From An Engineering Perspective: Capabilities And Limitations Of New Devices

Edward J. Delp
Video and Image Processing Laboratory
School of Electrical and Computer Engineering
School of Biomedical Engineering
Purdue University
West Lafayette, Indiana

ace@ecn.purdue.edu

https://engineering.purdue.edu/~ace/

http://www.tadaproject.org/

January 10, 2019
Food Image Analysis: The Big Data Problem You Can Eat!

Edward J. Delp  
Video and Image Processing Laboratory  
School of Electrical and Computer Engineering  
School of Biomedical Engineering  
Purdue University  
West Lafayette, Indiana

ace@ecn.purdue.edu

https://engineering.purdue.edu/~ace/

http://www.tadaproject.org/

January 10, 2019
Research Team

• Carol Boushey and Deborah Kerr

• Fengqing Zhu

• Students (not complete)
  – Sri Kalyan Yarlagadda, Yu Wang, Shaobo Fang, Chang Liu, Ziad Ahmad, TusaRebecca Schap, Ye He, Chang Xu, Marc Bosch Ruiz

• Acknowledge the National Institutes of Health, Curtin University, Lilly, CTSI, Purdue Research Foundation, The endowment of the Charles William Harrison Distinguished Professorship

• I have no financial interest
Measuring accurate dietary intake is considered to be an open research problem
Before Meal

Extract

Plate Waste

Food intake comparison

IlSI/VIPER January 10, 2019
Technology Assisted Dietary Assessment

Dietary intake provides some of the most valuable insights for mounting intervention programs for prevention. With the growing concern about overweight, the need to accurately measure diet becomes imperative. For example, assessment among adolescents is problematic as this group has irregular eating patterns and they have less enthusiasm for recording food intake. Preliminary studies among adolescents suggest that innovative use of technology may improve the accuracy of diet information from young people. Recognition of merging technology, e.g., higher resolution pictures, improved memory capacity, faster processors, allow these devices to process information not previously possible.

Our goal is to develop, implement, and evaluate a mobile device food record (mDiFR) that will translate to an accurate account of daily food and nutrient intake among adolescents and adults. Our first steps include further development of our pilot mobile computing device to include digital images, a nutrient database, and image processing for identification and quantification of food consumption. Mobile computing devices provide a unique vehicle for collecting dietary information that reduces burden on record keepers. Images of food can be marked with a variety of input methods that link the item for image processing and analysis to estimate the amount of food consumed.

This project is funded by the National Institutes of Health (NIH).

System Overview

Principal Investigators

Carol J. Boushey, Associate Professor, Epidemiology Program, University of Hawaii Cancer Center; and Adjunct Professor, Department of Nutrition Science, Purdue University

Edward J. Delp, The Charles William Harrison Distinguished Professor of Electrical and Computer Engineering and Professor of Biomedical Engineering, Purdue University

Peiminqing Maggie Zhu, Assistant Professor of Electrical and Computer Engineering, Purdue University

www.tadaproject.org
http://healthchat2.curtin.edu.au/

http://healthchat.curtin.edu.au/
TADA System Overview

1. **Image(s) + Metadata**
   - (Geolocation, Time, Barcode, Contextual Info)

2. **Volume Estimation**
   - Food Label Type
   - Segmented Image
   - Machine Learning
   - Context Processing
   - User Eating Patterns

3. **Communication Layer**
4. **Web Server**
5. **TADA Food Databases**
   - I-TADA
   - T-FNDDS
   - E-TADA
6. **Output**
   - Labeled Images With Food Type (e.g. Milk, Toast, Eggs)
7. **Wi-Fi/3G/4G/5G Network**

**User Feedback**
- Confirmation or Correction

**mFR**

**Cloud**

**www.tadaproject.org**
TADA Databases

Users/Patients/Participants

Nutritionist/Dietitian/Researchers/Participants

Communications Layer

Web Server

TADA Databases

I-TADA

T-FNDDS

E-TADA

Image Analysis and Volume Estimation

Wi-Fi/3G/4G/5G

Internet
Meal Image

Locate & Identify

Peach Ketchup Coke Milk

Hamburger French Fries Sugar Cookie

January 10, 2019
Example of Classification

Green – Correct classification
Red – Wrong classification
User Studies

• We have completed a total of more than 30 user studies
  – Free-living environment
  – More than 2400 participants
  – More than 170,000 food images acquired (with metadata)

• Each food image captures a real eating scene consists of multiple food items

• We have published more than 45 conference papers and 25 journal papers, 10 Ph.D. thesis, 2 patents and many software products designed, produced and published since 2008 (detailed information at: www.tadaproject.org)
### Summary of Studies: Completed or In Process
(not up to date)

**Age, y**
- 3
- 6
- 9
- 12
- 15
- 18
- 21
- 24
- 27
- 30
- 33
- 36
- 39
- 42
- 45
- 48
- 51
- 54
- 57
- 60
- 63
- 66

<table>
<thead>
<tr>
<th>Study number</th>
<th>Age range, y</th>
<th>Sample size</th>
<th>Supervised, EO</th>
<th>Community, # of days</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3-10</td>
<td>63</td>
<td>1-4</td>
<td>--</td>
</tr>
<tr>
<td>2</td>
<td>3-10</td>
<td>63</td>
<td>--</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>7-10</td>
<td>12</td>
<td>--</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>10-21</td>
<td>19</td>
<td>--</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>11-13</td>
<td>69</td>
<td>--</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>11-15</td>
<td>41</td>
<td>9-11</td>
<td>--</td>
</tr>
<tr>
<td>7</td>
<td>11-18</td>
<td>63</td>
<td>2</td>
<td>--</td>
</tr>
<tr>
<td>8</td>
<td>11-18</td>
<td>15</td>
<td>24 hr</td>
<td>--</td>
</tr>
<tr>
<td>9</td>
<td>12-30</td>
<td>59</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study number</th>
<th>Age range, y</th>
<th>Sample size</th>
<th>Supervised, EO</th>
<th>Community, # of days</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>14-21</td>
<td>18</td>
<td>--</td>
<td>56-58</td>
</tr>
<tr>
<td>11</td>
<td>18-30</td>
<td>247</td>
<td>--</td>
<td>8</td>
</tr>
<tr>
<td>12</td>
<td>18-55</td>
<td>20</td>
<td>--</td>
<td>8</td>
</tr>
<tr>
<td>13</td>
<td>20-70</td>
<td>77</td>
<td>--</td>
<td>8</td>
</tr>
<tr>
<td>14</td>
<td>21-65</td>
<td>57</td>
<td>2</td>
<td>--</td>
</tr>
<tr>
<td>15</td>
<td>21-65</td>
<td>45</td>
<td>--</td>
<td>7</td>
</tr>
<tr>
<td>16</td>
<td>21-65</td>
<td>22</td>
<td>--</td>
<td>2</td>
</tr>
<tr>
<td>ALL</td>
<td>3-70</td>
<td>890</td>
<td>1-11</td>
<td>2-58</td>
</tr>
</tbody>
</table>

**KEY:** y=years, EO=eating occasions
<table>
<thead>
<tr>
<th>Community Dwelling</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Day 0</strong></td>
</tr>
<tr>
<td>Thu</td>
</tr>
<tr>
<td>n = 46</td>
</tr>
</tbody>
</table>
University of Hawaii: Study Eating Behaviors of Children In Guam

How old is this participant?

9 years old!
TADA Color Fiducial Marker

- TADA color fiducial marker is a reference for food classification and portion estimation:
  - Geometric reference
  - Color reference
  - Image quality reference

- Real time image quality check is implemented on the mobile phone

TADA color fiducial marker
Mobile Food Record (mFR)

Apple (iOSX) – iPhone, iPod, iPad

Android (4.3 and above) – phones and tablets

API and SDK are available for various parts of the TADA architecture
Image Acquisition

- Angle information is obtained from the phone

- Colors along with words assist the user in taking an image at the preferred angle
Image Quality Checking

• Check for:
  – Presence of fiducial marker (checkerboard)
  –Blurry image
Review

• The user can select an eating occasion from the list to review
• The before eating image is then displayed in landscape view with food labels on it
Review

• Different colors help the user in identifying bubble-pin correspondence

• The green color is reserved for confirmed labels

• Zoom-in by pinching the screen to have a better view
TADA Image Analysis
Early Approach

1. Color
2. Texture
3. Local

Food Image → Segmentation → Feature Extraction

Segmentation Refinement

Vocabulary Tree + KNN
TADA Image Analysis System

Training Dataset → Input Image → Segmentation

- Global Feature Channels
  - Global Feature Channel 1
  - Global Feature Channel LG
  - Global Feature Channel LL

- Local Feature Channels
  - Point of Interest Detector
  - Local Feature Channel 1
  - Local Feature Channel LL

- Vocabularies
  - Local Feature Channel 1
  - Local Feature Channel LL

- Segmentation
  - Late Decision Fusion
  - Early Decision Fusion

- Output
  - Portion Estimation
  - Context Refinement

Segmentation Improvement → YES/NO
Segmentation

• Multiscale normalized cut

Increasing Number of Segments
TADA Deep Learning
Food Image Processing

Original Image

Food Localization

Food Classification

Food Segmentation

milk  sausage  pancake

milk  sausage  pancake
Food Localization

• Method
  – Faster RCNN + ResNet-101
  – Deformable Convolution Layer (DCL)
Food Classification

• Method
  – Inception-v3 Network

![Early Inception Module Diagram]
Classification Accuracy of Foods
The Use of Contextual Information

- Time-based food consumption frequency
- Food co-occurrence patterns
- Date, time, place, dietary habits (patterns), work/sleep patterns
Food Co-Occurrence Patterns

• The likelihood of food combinations --- their mutual probability of existing together in a single eating occasion

• A post-processing stage to promote agreement between the segment labels
Food Co-Occurrence Patterns

- Build a fully connected undirected graph between all segments
- Adjust the probability of each node by its association with all other nodes

\[
p'(f_k|S_n) = \frac{p(f_k|S_n)A(f, S)}{Z(\phi, S_1, ..., S_N)}
\]

\[
A(f, S) = \exp(\sum_{i=0}^{4} \sum_{j=1, j \neq n}^{N} \phi(f_{k,n}, f_{i,j})p(f_i|S_j))
\]

where \(p(f_k|S_n)\) is the probability of the food label \(f_k\) for segment \(S_n\)

\(\phi(f_{k,n}, f_{i,j})\) is the co-occurrence probability between two nodes

\(Z(.)\) is the normalization constant obtained by summing the numerator over all nodes in the same clique
Food Co-Occurrence
Temporal Context

Recursive Bayesian Model of Food Consumption Frequency

• Let $p_{\lambda_i}(x^k)$ be the probability density function (PDF) representing $S_j$ consumes $\lambda_i$ on the $k^{th}$ day, and $z^k$ be the observation whether $S_j$ consumes $\lambda_i$ on the $k^{th}$ day.

• $z^k$ is obtained from the user feedback in the TADA system.

• Posteriori update:

$$p_{\lambda_i}(x^k | z^{1:k}) = \frac{p_{\lambda_i}(x^k | z^k)p_{\lambda_i}(x^k | z^{1:k-1})}{p_{\lambda_i}(z^k | z^{1:k-1})} = \frac{\text{likelihood} \times \text{prior}}{\text{normalization}}$$

• On the $k+1^{th}$ day, $P_{\lambda_i}$ is computed as $P_{\lambda_i} = \arg \max p_{\lambda_i}(x^k|z^{1:k})$
The Use of Temporal Context

• Selected participants with similar food consumption patterns were used to build personalized eating datasets for a month

• Three separate datasets (i.e. dataset 1, 2 and 3) with a total of 358 food images from a free-living study

<table>
<thead>
<tr>
<th>statistics</th>
<th>user ID</th>
<th>with context</th>
<th>without context</th>
</tr>
</thead>
<tbody>
<tr>
<td>average daily</td>
<td>user1</td>
<td>61.88</td>
<td>53.23</td>
</tr>
<tr>
<td>classification</td>
<td>user2</td>
<td>65.25</td>
<td>62.90</td>
</tr>
<tr>
<td>accuracy(%)</td>
<td>user3</td>
<td>59.69</td>
<td>53.28</td>
</tr>
<tr>
<td>average daily</td>
<td>user1</td>
<td>18.45</td>
<td></td>
</tr>
<tr>
<td>accuracy</td>
<td>user2</td>
<td>3.85</td>
<td></td>
</tr>
<tr>
<td>improvement(%)</td>
<td>user3</td>
<td>12.39</td>
<td></td>
</tr>
</tbody>
</table>
Food Portion Estimation: Single-View

- To reduce a user’s burden, our work has focused on the use of a single image to estimate food portion.

- Food portion estimation based on a single-view is an ill-posed inverse problem.
  - Most 3D information has been lost during projection process.

- We use pre-defined geometric models to estimate food portion size.
## Food Volume Estimation

**Input datasets**

- **Meal Image**
- **Segmented Images**

**Results**

- **Best-Fit Geometric Model**
  - Cylinder, Box, Prism
- **Estimated volumes**

**Feature Points**

<table>
<thead>
<tr>
<th>Food Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>11112110</td>
<td>Milk</td>
</tr>
<tr>
<td>53105500</td>
<td>Chocolate cake</td>
</tr>
</tbody>
</table>

**Example**

- **Feature Points**
  - [Image of feature points]

- **Volume Estimation**
  - Feature point extraction
  - 3D volume reconstruction

- **Nutrient Computation**
  - [Table of nutrient facts]

- **Camera Calibration**
  - [Images of calibrated camera parameters]

- **Shape templates search**
  - Lookup Table

- **Best fit template**
  - Food code (FNDDS)
Generate or Pre-Define Food Shapes

<table>
<thead>
<tr>
<th>Shape</th>
<th>Example Food Type</th>
<th>Dimension Parameters</th>
<th>Locator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cylinder</td>
<td>Orange juice, Milk</td>
<td>Radius, Height</td>
<td></td>
</tr>
<tr>
<td>Sphere</td>
<td>Apple, Orange</td>
<td>Radius</td>
<td></td>
</tr>
<tr>
<td>Square Box</td>
<td>Chocolate Cake, Brownie</td>
<td>Width, Length, Height, Rotation Angle</td>
<td></td>
</tr>
<tr>
<td>Slice of Cone/Slice of Sphere</td>
<td>Spaghetti, Ice Cream</td>
<td>Top Radius, Bottom Radius, Height</td>
<td></td>
</tr>
<tr>
<td>Prism</td>
<td>Bread, Scrambled Eggs</td>
<td>Area, Height</td>
<td></td>
</tr>
<tr>
<td>Irregular Shape</td>
<td>Banana, Pear</td>
<td>Scale X, Scale Y, Scale Z, (Rotation Angle)</td>
<td></td>
</tr>
</tbody>
</table>
Food Items

1. 2% Milk
2. Sausage links
3. Scrambled eggs
4. Toast
5. Garlic bread
6. Chocolate cake w/ icing
7. Sugar cookie
8. Spaghetti w/ sauce, cheese
9. Orange juice
10. Peach slices
11. Pear, canned halves
12. French fries
13. Ketchup
14. Lettuce (salad)
15. Margarine
16. French dressing
17. Strawberry jam
18. Coke
19. Cheeseburger sandwich

Fang et al. ISM 2015
Food Density

\[
\text{Density } \left( \frac{\text{g}}{\text{cm}^3} \right) = \frac{\text{Weight (g)}}{\text{Volume (cm}^3\text{)}}
\]

- **True density**: density of pure substance or material calculated from its component densities
- **Apparent density**: density of a particle including all pores
- **Bulk density**: density when particles are packed or stacked in bulk including void spaces
The Estimated Energy For A Meal

• Three sample meals

Combination type A: average ratio of estimated energy to ground truth: **1.01**

Combination type B: average ratio of estimated energy to ground truth: **0.97**

Combination Type C: average ratio of estimated energy to ground truth: **1.06**
Ongoing Work

• VIPER-FoodNet (VFN) dataset: 2.5M+ food images from the net verified by crowdsourcing

• Passive vs Active Approaches

• Semantic segmentation and transfer learning

• Depth prediction from a single image (one shot learning)

• From supervised learning to unsupervised learning (handle large number of classes)
Ongoing Work

• Use of Generative Adversarial Networks (GANs) for both food recognition and portion size estimation

• The extraction of dietary patterns

• Combining activity measures with food images and investigating new types of dietary patterns
Takeways

• Modern machine learning methods will change how we can do dietary assessment

• Active vs. Passive approaches
Food Image Analysis:
The Big Data Problem You Can Eat!

Edward J. Delp
Video and Image Processing Laboratory
School of Electrical and Computer Engineering
School of Biomedical Engineering
Purdue University
West Lafayette, Indiana

ace@ecn.purdue.edu

https://engineering.purdue.edu/~ace/

http://www.tadaproject.org/

January 10, 2019