MT-Diet Demo: Demonstration of Automated Smartphone Based Diet Assessment System

Junghyo Lee, Ayan Banerjee, and Sandeep K. S. Gupta
IMPACT Lab, CIDSE, Arizona State University, Tempe, AZ
Email: {jlee375,abanerj3,sandeep.gupta}@asu.edu

***Abstract*—Background: According to several recent research results [1]–[4], obesity can increase the risk of many diseases such as diabetes, chronic kidney disease, metabolic disease, cardiovascular disease, etc. To prevent and treat the obesity efficiently and effectively, diet monitoring is an important factor. Purpose: Manual self-monitoring techniques for diet suffer from drawbacks such as low adherence, underreporting, and recall error [5]–[7]. Camera based applications that automatically extract type and quantity of food from an image of the food plate can potentially improve adherence and accuracy. However, state-of-the-art systems [8] have fairly low accuracy for identifying cooked food (only 63%) and are not fully automatic. To overcome these drawbacks such as low adherence, underreporting, recall error, low accuracy, and semi-automatedness, we introduce *MT-Diet*, a fully automated diet assessment system. It can identify cooked food with an accuracy of 88.93%. This is a significant improvement (over 20%) from the current state-of-the art system. Method: MT-Diet is a smartphone-based system that interfaces a thermal sensor with a smartphone. Using this system a user can take both thermal and visual images of her food plate with just one click. We used a database of 80 frozen meals which contain several different types of foods so that the actual total number of our food database 244 and the database has 33 different types of foods. By using the database, we demonstrate two core components: a) food segmentation, separating food items from the plate and recognizing multiple food items as a single food item, and b) food identification, determining the type of foods. Result: MT-Diet food segmentation methodology is fully automatic and requires no user input as opposed to recent works, the accuracy of separating food parts from the plate was 97.5%. The accuracy of food identification using Support Vector Machine with Radial Basis Function kernel based on color, texture, and histogram of oriented gradients features is 88.5%. Conclusion: We suggest a new and novel approach for diet assessment, MT-Diet. Our approach can potentially be an inexpensive, real time for the feedback on calorie intake, easy-to-use, privacy preservation, personalization based on eating habits of individuals, and fully automated diet monitoring system. The tool can also be used to conduct clinical studies to develop models of meal patterns that can be incorporated to design better artificial pancreas.**

I. Introduction

Accurate assessment of dietary intake is important for: i) analyzing the relationship between caloric intake and health outcomes such as obesity, ii) evaluating outcomes of dietary change interventions, iii) ensuring compliance with dietary recommendations, iv) self-motivated diet control, or v) identifying factors that induce change in dietary intake so that they can be used in targeted interventions. Recent studies have reported that smartphone users find image based diet monitoring systems easy to use and more usable than text based records [9]. Another study on a medium size cohort of adults aged 18 to 24 years has reported that nearly 87 % of users feel that they will not be opposed using a food image based diet monitoring mobile application for long term [10]. However, feedback from a study on the usability of MyFitnessPal [11], an image based diet monitoring application indicate that users would prefer more information related to type of food and calorie intake from just an image upload. According to recent surveys [12]–[14], image based applications that automatically extract type and quantity of food from an image of the food plate have good accuracy for identifying fresh food. However, a larger share of daily calorie intake comes from hot cooked food for which these systems have fairly low identification accuracy (nearly 63% accuracy at best). This is because hot/warm foods also tend to be mixed dishes (e.g., lasagna), which are difficult to assess using color images. We demonstrate MT-Diet, a smartphone based automated cooked food identification system that can determine the type of cooked food on plate with nearly 90% accuracy.

MT-Diet interfaces a thermal sensor with a smartphone, so that a user can take both thermal and visual images of his/her food plate with just one click. The system uses thermal maps of a food plate to increase accuracy of extraction and segmentation of food parts, combines thermal and visual images to improve accuracy in the detection of cooked food. Preliminary testing results show that *MT-Diet* can determine the type of food consumed with an average accuracy of 90%, which is a significant improvement from the current state-of-the-art. *MT-Diet* is implemented as a part of the bHealthy application suite for behavioral healthy monitoring [15].

This project is partially funded by NSF IIS 1116385 and NIBIB EB019202.

II. Demonstration Setup and Plan

MT-Diet application

Requirements: a) cooked foods, b) a smartphone built in a camera, c) a seek thermal sensor [16], and d) two small caps filled with cold water.

Inputs: a) a plate full of hot food, b) color image from smartphone camera, and c) thermal image from infrared camera.

Outputs: a) Segmented food images, b) Food type in plate, and c) Nutrition information into USDA website.

Platform: a) a smartphone interfaced with thermal camera and b) reliable connection of the smartphone with a cloud server.

Assumptions:
a) Food temperature $\gg$ Plate temperature;
b) Plate temperature > Background temperature;
c) The plate is not overflowing with food; and
d) A database of food items is prepared offline and available to *MT-Diet*.

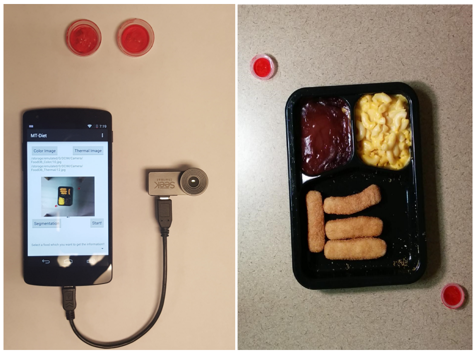

Fig. 1. Actual requirements: Left are Nexus 5, Seek thermal sensor, and two small caps. Right is a cooked frozen food.

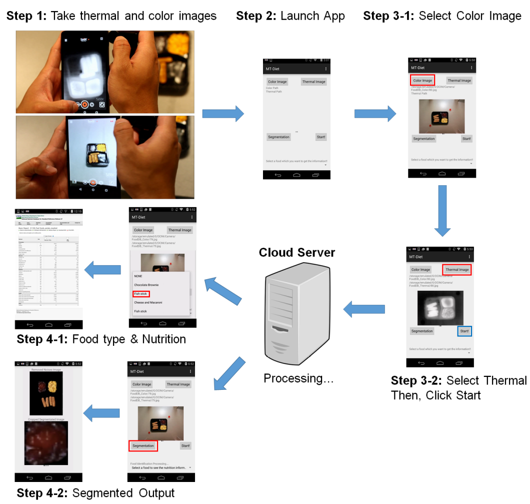

Fig. 2. *MT-Diet* application steps for the demonstration.

Before launching the application, we need to prepare cooked foods, a smartphone with built-in a camera, and a Seek thermal sensor [16] interfaced with the smartphone such as Fig. 1. Frozen foods are used since it is affordable and easy to get. These foods are defrosted for 15 minutes using a microwave. Then, we take the image of cooked food using the smartphone camera and the thermal camera. The Seek camera is connected to the Nexus 5 phone using a micro-USB chord, due to which, the resulting images have different angles and distances. Hence, we put two small caps filled with cold water at two diagonally opposite ends of the plate for the calibration purpose. After taking these two images by these cameras, we are ready to launch the *MT-Diet* application. The application needs both a color image and a thermal image of the food as inputs. The application then provides the user with (a) segmented food images (removed plate and background), (b) food type in plate, and (c) each food's nutrition information into USDA (United States Department of Agriculture) website.

The demonstration consists of following steps: **(a) The user selects from color image and thermal image by clicking two buttons: Color Image and Thermal Image such as Step 3-1 and 3-2 in Fig. 2** . After clicking these buttons, the application opens the gallery and the users can find the images that they took. The application displays the images as well as their absolute paths, the user can now visually verify whether the selected images are correct. **(b) The user clicks the start button to send the images to the cloud server**. The cloud server connects the application using the socket communication. Fig. 3 illustrates the connection between the server and the application. As seen in the figure, the application sends the two images to the server. **(c) The server processes the food recognition included the food segmentation and food identification**. The application produces accurate food segmentation without needing to to ask the user by combining three algorithms: Dynamic Thermal Threshold (DTT), Hierarchical Image Segmentation (HIS) [17], and Grabcut [18]. The approach and algorithms in the detail are provided in a paper [19] Also, the execution time of the food identification is fast since the server already has the trained parameters for SVM classification. **(d) The server sends the segmented images and food types to the application.** After receiving the food types, the application can find the USDA database links of the food type by matching with the USDA database.

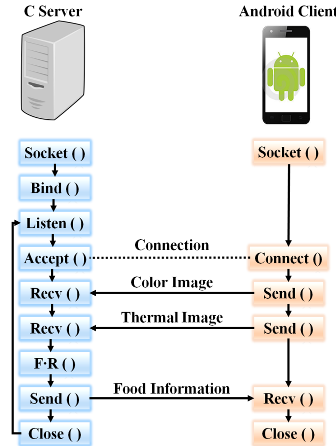

Fig. 3. Flow Chart of MT-Diet by Socket.

The food types are displayed as a spinner button (Step 4-1 in Fig. 2) and the segmented images clicking the Segmentation button (Step 4-2 in Fig. 2). After looking at the segmented images, the user can judge each segmented image quality, which is critical for the identification accuracy. Moreover, the user can access the USDA database about the food by clicking the food type spinner button. Therefore, by taking two images, the user can obtain the accurate food information automatically. The application is envisioned to be inexpensive, easy-to-use, and privacy preserving. In addition the application can provide real-time feedback on calorie intake and can be personalized based on eating habits of individuals.

A video of the demonstration is available in youtube [20].

III. Special requirements

For the demo, we need (a) Microwave with power outlet to cook the frozen foods and (b) WiFi Network to connect to the cloud server.

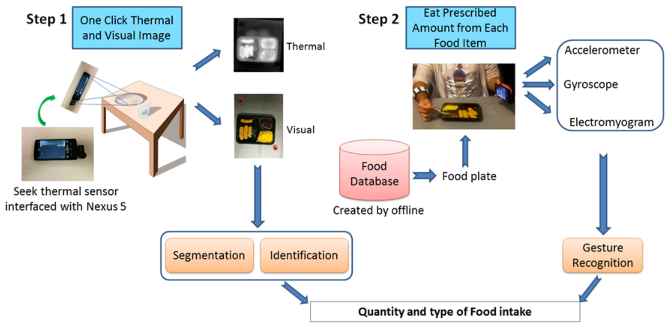

Fig. 4. System Architecture of the extended *MT-Diet* food monitoring system including the measuring food consumption.

IV. Discussion and Future Work

Since this demo is in an initial phase, there are some insufficient implementations so in this section, we will discuss the future phases. The first extension we propose is a technique such that the two bottle caps that were used for calibration are no longer needed. Using the bottle caps limits the usage of the application so we suggest an extended implementation. In the initial implementation, Nexus 5 needs the micro USB wire because the direction of the thermal camera lens is opposite to the direction of a built-in Nexus 5 camera. The micro USB wire is an obstacle for calibration because it generates different angles and distance between the thermal image and color image. So we will change the smartphone to LG G2 which does not need the micro USB wire to connect the thermal camera. Then, we find the fixed different angle and distance the thermal image and color image by executing only one calibration task. Therefore, the small caps are not required for the calibration process.

In addition, the the application requires high computation time (almost 100 seconds for the segmentation) even when the cloud server is employed. So, we consider parallel computation as a potential solution of the issue. Moreover, if the user is in a situation where there is no network connection, the application cannot work, thus we pursuit an approach so that the application will work in a non-network environment. To deal with both issues, we consider changing the programming language from Matlab to C++ because C++ not only supports cross compiling with Android but also can be extended to OpenCL for smartphone GPU implementation.

Food identification is a important task for diet monitoring, but estimating food consumption is also a critical issue. Our idea for the measuring the user's food consumption is to recognize the user hand movements using the wrist sensors as seen in Step 2 in Fig. 4. To measure the food consumption using the sensor, we need to consider three problems: a) identifying utensils such as fork, spoon, chopstick, and knife b) mapping food image location and actual food location and c) identifying and counting user's eating motion.

V. Conclusion

Automated diet monitoring and caloric intake prediction is an effective intervention for chronic diseases such as cardiac problems, obesity and diabetes that affect more than one-third of US adults with a combined estimated economic $392 Billion cost. In this project we evaluate the accuracy and usability of *MT-Diet*, a cost effective smartphone based automated diet monitoring application that uses the image of a food plate in both the thermal and visual spectra to identify food type. Such an easy-to-use and cost effective solution to real time diet monitoring can potentially achieve higher adherence to interventions which may in turn lead to beneficial health impacts such as effective weight reduction.

References

[1] C. Zhang, K. M. Rexrode, R. M. van Dam, T. Y. Li, and F. B. Hu, "Abdominal obesity and the risk of all-cause, cardiovascular, and cancer mortality sixteen years of follow-up in us women," *Circulation*, vol. 117, no. 13, pp. 1658–1667, 2008.

[2] M. T. Hamilton, D. G. Hamilton, and T. W. Zderic, "Role of low energy expenditure and sitting in obesity, metabolic syndrome, type 2 diabetes, and cardiovascular disease," *Diabetes*, vol. 56, no. 11, pp. 2655–2667, 2007.

[3] H. Bays and C. Ballantyne, "Adiposopathy: why do adiposity and obesity cause metabolic disease?" 2006.

[4] J. E. Hall, J. R. Henegar, T. M. Dwyer, J. Liu, A. A. da Silva, J. J. Kuo, and L. Tallam, "Is obesity a major cause of chronic kidney disease?" *Advances in renal replacement therapy*, vol. 11, no. 1, pp. 41–54, 2004.

[5] T. L. Burrows, R. J. Martin, and C. E. Collins, "A systematic review of the validity of dietary assessment methods in children when compared with the method of doubly labeled water," *Journal of the American Dietetic Association*, vol. 110, no. 10, pp. 1501–1510, 2010.

[6] C. M. Champagne, G. A. Bray, A. A. Kurtz, J. B. R. Monteiro, E. Tucker, J. Volaufova, and J. P. Delany, "Energy intake and energy expenditure: a controlled study comparing dietitians and non-dietitians," *Journal of the American Dietetic Association*, vol. 102, no. 10, pp. 1428–1432, 2002.

[7] J. R. Hebert, C. B. Ebbeling, C. E. Matthews, T. G. Hurley, M. Yunsheng, S. Druker, and L. Clemow, "Systematic errors in middle-aged women's estimates of energy intake: comparing three self-report measures to total energy expenditure from doubly labeled water," *Annals of epidemiology*, vol. 12, no. 8, pp. 577–586, 2002.

[8] M.-Y. Chen, Y.-H. Yang, C.-J. Ho, S.-H. Wang, S.-M. Liu, E. Chang, C.-H. Yeh, and M. Ouhyoung, "Automatic chinese food identification and quantity estimation," in *SIGGRAPH Asia*. ACM, 2012, p. 29.

[9] K. Aizawa, K. Maeda, M. Ogawa, Y. Sato, M. Kasamatsu, K. Waki, and H. Takimoto, "Comparative study of the routine daily usability of foodlog a smartphone-based food recording tool assisted by image retrieval," *Journal of diabetes science and technology*, vol. 8, no. 2, pp. 203–208, 2014.

[10] S. Sharma and S. Fulton, "Diet-induced obesity promotes depressive-like behaviour that is associated with neural adaptations in brain reward circuitry," *International journal of obesity*, vol. 37, no. 3, pp. 382–389, 2013.

[11] [Online]. Available: https://www.myfitnesspal.com/.

[12] P. Pouladzadeh, P. Kuhad, S. V. B. Peddi, A. Yassine, and S. Shirmohammadi, "Mobile cloud based food calorie measurement," in *Multimedia and Expo Workshops (ICMEW)*. IEEE, 2014, pp. 1–6.

[13] P. Pouladzadeh, S. Shirmohammadi, A. Bakirov, A. Bulut, and A. Yassine, "Cloud-based svm for food categorization," *Multimedia Tools and Applications*, pp. 1–18, 2014.

[14] P. Pouladzadeh, S. Shirmohammadi, and A. Yassine, "Using graph cut segmentation for food calorie measurement," in *Medical Measurements and Applications (MeMeA)*. IEEE, 2014, pp. 1–6.

[15] J. Milazzo, P. Bagade, A. Banerjee, and S. K. S. Gupta, "bhealthy: A physiological feedback-based mobile wellness application suite," in *Proceedings of the 4th Conference on Wireless Health*, ser. WH '13. New York, NY, USA: ACM, 2013, pp. 14:1–14:2.

[16] Accessed: 2015-07-10. [Online]. Available: http://www.thermal.com/.

[17] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 5, pp. 898–916, 2011.

[18] C. Rother, V. Kolmogorov, and A. Blake, "Grabcut: Interactive foreground extraction using iterated graph cuts," *ACM Transactions on Graphics (TOG)*, vol. 23, no. 3, pp. 309–314, 2004.

[19] J. Lee, A. Banerjee, and S. K. S. Gupta, "Mt-diet: Automated smartphone based diet assessment with infrared images," in *Pervasive Computing and Communications (PerCom)*. IEEE, 2016.

[20] [Online]. Available: https://www.youtube.com/watch?v=En8iyJ5JSI4.