# Distinguishing Nigerian Food Items and Calorie Content with Hyperspectral Imaging

Xinzuo Wang[1], Neda Rohani[1], Adwaiy Manerikar[1], Aggelos Katsagellos[1], Oliver Cossairt[1], and Nabil Alshurafa[1,2]

[1] Department of EECS, Northwestern University, Evanston, IL, USA
[2] Feinberg School of Medicine, Northwestern University, Chicago, IL, USA
`{XinzuoWang2018,nedarohani,AdwaiyManerikar2019}@u.northwestern.edu`,
`{oliver.cossairt,nabil}@northwestern.edu}`
`aggk@eecs.northwestern.edu`

**Abstract.** Identifying food types consumed and their calorie composition is one of the central tasks of dietary assessment. Traditional automated image processing methods learn to map images to an existing food database with known caloric composition. However, even when the correct food type is identified, caloric makeup can vary depending on its ingredients, and using true-color images proves insufficient to distinguish within food type variability. In this paper, we show that hyperspectral imaging provides useful information and promise in distinguishing caloric composition within the same food type. We collect data using a hyperspectral camera from Nigerian foods cooked with varying degrees of fat content, and capture images under different intensities of light. We apply Principle Component Analysis (PCA) to reduce the dimensionality, and train a Support Vector Machine (SVM) classifier using a Radial Basis Function kernel and show that applying this technique on hyperspectral images can more readily distinguish calorie composition. Furthermore, compared with methods that only use true-color based features, our method shows that a classifier trained using features from hyperspectral images is significantly more predictive of within-food caloric content, and by fusing results from two classifiers trained separately using hyperspectral and RGB imagery we obtain the greatest predictive power.

**Keywords:** hyperspectral imaging, food identification, calorie detection

## 1 Introduction

Food crisis and undernutrition have been critical issues in many low and middle income countries. One promising way to address this problem is to provide targeted delivery of health and supplementary nutrition (like the Plumpy'Nut project in East Africa), food fortification, and empowering local villages to grow nutritious foods.

However, aside from daily food diaries and reports from local health workers, there is no reliable automated method to identify foods consumed and their

caloric content. A method is needed to detect - for every food consumed - a unique food signature that is invariant to light, heat, and slight variations in visual appearance.

Despite several efforts in the field of image recognition, conventional imaging technologies that only acquire morphology are not adequate for the accurate detection and assessment of intrinsic properties of a food. Spectroscopic imaging covering the visible and near-IR spectrum (from 400nm to 1100nm) can help identify unique spectral features that readily discriminate between foods. In this paper, through feature extraction and classification on hyperspectral images of real food, we show that hyperspectral imaging provides useful information and promise in distinguishing caloric composition within the same food type.

## 2  Related Works

Throughout the past decade, food images have been widely studied. For food image segmentation, Zhu et al.[13] employs connected component analysis and normalized cuts. Anthimopoulos et al. [1] uses mean-shift algorithm. Kong et al.[6] extracts Scale Invariant Feature Transform (SIFT) points. Matsuda et al. [8] provides a method to segment food images containing multiple food items by using Felzenszwalb's deformable part model (DPM) and JSEG region segmentation. For food feature extraction, Gabor features [13], simple color and text features [12] have been explored on food images. In Dehais et al.[3], SIFT and Speeded Up Robust Features (SURF) detectors are used.

Meanwhile, hyperspectral imaging has developed wide uses for identifying the chemical and physical properties of food. Some of these uses include predicting color [9], detecting damage [4], and analyzing quality attributes [5]. Also many works have been focused on bruises detection [7, 2], and food quality classification [10] by using PCA.

However, for food image recognition, while these methods have been successful in various data sets, they are all based on true color images which only focus on RGB channels. In this paper, we argue that RGB channels can be insufficient in many natural settings where foods and dishes are more complex (i.e. food in restaurants). To solve this problem, we take advantage of hyperspectral images which contain information from multiple spectra. In hyperspectral image analysis, only a few methods have been proposed to detect calorie contents of real-world foods. This work provides a method to distinguish calorie content of food types using hyperspectral imaging.

## 3  Methods

In this section we introduce our hyperspectral image processing technique. To begin, we first introduce our data (foods cooked), then our image acquisition system, followed by our food type identification and within-food calorie discriminant techniques.

### 3.1  Food Samples and the Hyperspectral Imaging System

To simulate real-world foods in low income countries, we prepared food samples using traditional Nigerian recipes. Three separate dishes were made with

three different levels of fat content, for a total of nine samples. The three dishes were white rice, chicken stew, and spiced yams. The fat content was adjusted by using various quantities of oil during cooking. The low-fat rice was prepared without butter while the medium- and high-fat rice dishes were prepared using one and four tablespoons of butter, respectively. Similarly, for the low-fat yam and stew no oil was used. One tablespoon of oil was used for the medium fat dishes, and four tablespoons of oil were used for the high fat dishes. The nutritional information for each dish was calculated based on the nutritional value of the ingredients. The nutritional information can be viewed in Table 3.1. During

**Table 1.** Nutritional information for food samples.

| Dish | Rice | | | Yams | | | Stew | | |
|---|---|---|---|---|---|---|---|---|---|
| | Low fat | Med fat | High fat | Low fat | Med fat | High fat | Low fat | Med fat | High fat |
| Total fat (kcal) | 204 | 306 | 612 | 249 | 351 | 657 | 577 | 697 | 1057 |

the measurement process, the system was shielded from outside light in order to decrease the interference of ambient light. The size of the recorded images is 240 wavebands with a resolution of 640 x 244 pixels. Image sizes varied slightly depending on the location of the camera at the beginning of each image acquisition. Data was collected for each pixel in the wavelength range of 393 nm to 892 nm.

### 3.2 Hyperspectral Image Acquisition

A total number of 30 images were taken. For each dish, three images were taken: one in each light setting. This led to a total of 27 images (9 x 3), followed by three extra reference images taken in between light adjustments to calibrate the system to the newer lighting settings. The images were acquired using a laboratory hyperspectral system. Only halogen lights were used, and all other light sources in the room were turned off. During image acquisition, 100 grams of each sample were placed in straw fiber bowls.

Due to the variant intensity of the light source, calibration with light and dark references was necessary in order to obtain accurate hyperspectral images. The dark reference was used in order to remove dark current effects. This image was collected by placing a black cap over the camera lens and turning off all lights in the room. The light reference was obtained by taking an image of a reflective white sheet of paper. The corrected image can be calculated by $R = R_o - R_d/R_r - R_d$ where $R_o$ is the acquired original hyperspectral image, $R_r$ is the white reference image, $R_d$ is the dark image.

This procedure was repeated twice more, and each time the two light sources were moved to a greater distance from the sample. Each time, the light source was adjusted, the images used for references had to be retaken in order to maintain an accurate reference measurement. The exposure time and camera speed remained the same for all images.

### 3.3 Feature Extraction for Food Item Detection

To identify between food types, we perform preprocessing, feature extraction, and classification, each section is described below:

**Preprocessing** The whole preprocessing procedure includes three steps: data cleaning, dimension reduction, and patch selection. During data cleaning we remove noisy bands between 360 nm to 480 nm (first 30 channels), so the wavelength range used for feature extraction is between 480 nm and 892 nm. The resulting data set includes 30 images (644 x 244 pixel resolution each) with 210 wavebands.

To further reduce the dimension of the data, we merge 210 wavebands into seven larger wavebands. That is, for each pixel, we calculate the mean for every 30 wavebands. Therefore, each pixel is represented by a seven-dimensional vector. Finally, we select 40 30 x 30 pixels patches from each image at each waveband to expand our data set and also enrich the training of our model.

**Feature extraction** The left side of Fig. 3.3 shows different food items we used for feature extraction. From the image, we can easily see that the visual appearance of different food items is largely different, while the same food items almost look alike and are homogeneous. This reminds us that simple statistical features such as mean and standard deviation from the visible light spectrum are enough for food type detection, particularly with a low selection of food types.
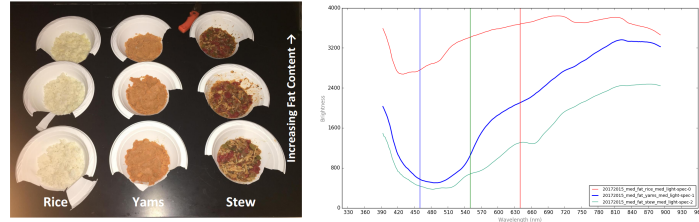


**Fig. 1.** Left side: Picture of our food samples. Right side: Average spectrum for three kinds of food items under the same light condition and fat content.

**Classification** For food item detection, we extract the mean, standard deviation, max and min from the first three wavebands for a total of 12 features. We then build a random forest classifier using these features to classify three kinds of food items.

### 3.4 Feature Extraction for Calorie Content Detection

After classifying the food type, we need to classify the food calorie content. The variation in calorie content is represented in three labels of low, medium, and high fat. As mentioned before, the images have been captured under different light conditions: high, medium, and low light. Images captured with different fat content and different light conditions of stew and yams have very similar spectra

which makes distinguishing between them very challenging. In Fig. 2, the spectra of stew with medium and high fat in three different light conditions are provided and as can be observed they have very similar shapes and intensities. For fat
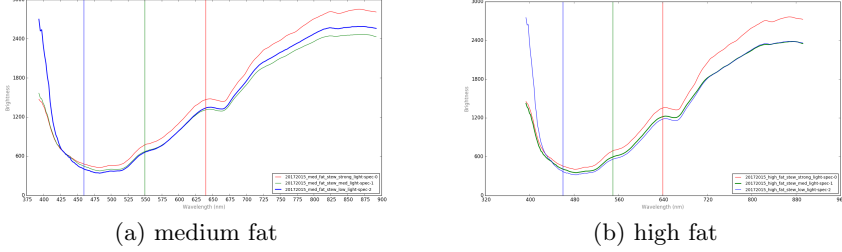


(a) medium fat                                        (b) high fat

**Fig. 2.** Images of stews with medium and high fat content captured under different light conditions. Even though the system was calibrated between each light condition, there are still differences in the spectra based on the intensity of light.

content classification, for each food, we use nine images: three images under different light conditions for each fat condition. Before classifying the calorie content we apply different preprocessing and feature extraction described below.

**Preprocessing** First, we crop each image and select a rectangle of the image that contains the food. Then, we remove the first 30 noisy bands. To have a more meaningful data representation in a lower dimensional space, we apply PCA and select the first k-components which accounts for much of the variance in the data (accounting for 98% variance of the information). We selected 16, 51 and 43 PCs for rice, yams and stew, respectively.

**Feature Extraction** After preprocessing the images, we divide the data into train and test sets. The images in two different light conditions (e.g. low and high) are considered as training images and the images in another light condition (e.g. medium) are used as test images. Now that we have training and test images, we extract patches of size 30x30xk from each image and use the "mean" spectrum over the patch pixels as the new feature vector (size k).

## 4    Experiments and Results

In this section we apply classification methods to the food data sets to evaluate selected features.

### 4.1    Food Item Classification

The random forest yielded best performance with 10 trees and a maximum depth of 10. The confusion matrix for three labels is shown in Table 2. Results show that the classifier achieves more than 98% accuracy.

**Table 2.** Confusion matrix and F-measure for food item classification.

| Classified as | Yams | Rice | Stew |
|---|---|---|---|
| | (F-measure: 0.976) | (F-measure: 0.997) | (F-measure: 0.977) |
| Yams | 346 | 0 | 14 |
| Rice | 0 | 359 | 1 |
| Stew | 3 | 1 | 356 |

## 4.2 Food Calorie Content Classification

For the fat content classification, we use Radial Basis Function (RBF) kernel SVM which has been used widely for classification problems. In RBF, the kernel function is: $K(x_i, x_j) = \exp(-\gamma \| x_i - x_j \|_2^2)$. The optimal values of hyper parameters of RBF classifier are obtained by using n-fold cross validation. The optimal $\gamma$ values have been found as $1.1e-2$, $1.9e-2$ and $2.3e-2$, for rice, yams and stew respectively. Then, the classifiers are trained with the optimal hyper parameters and tested on the test dataset.

**Table 3.** Results for food calorie content identification using RGB and hyper-spectral images.

| | Rice | | | Yams | | | Stew | | |
|---|---|---|---|---|---|---|---|---|---|
| RGB images | (Accuracy: 91.66%) | | | (Accuracy: 65.83%) | | | (Accuracy: 54.44%) | | |
| Classified as | Low fat | Med fat | High fat | Low fat | Med fat | High fat | Low fat | Med fat | High fat |
| Low fat | 60 | 0 | 0 | 33 | 7 | 0 | 43 | 2 | 15 |
| Med fat | 14 | 45 | 1 | 17 | 19 | 4 | 31 | 20 | 9 |
| High fat | 0 | 0 | 60 | 6 | 7 | 27 | 25 | 13 | 22 |
| Precision | 0.81 | 1 | 0.98 | 0.59 | 0.58 | 0.87 | 0.43 | 0.57 | 0.47 |
| Recall | 1 | 0.75 | 1 | 0.82 | 0.47 | 0.67 | 0.71 | 0.33 | 0.36 |
| | Rice | | | Yams | | | Stew | | |
| hyper-spectral images | (Accuracy: 99.44%) | | | (Accuracy: 81.66%) | | | (Accuracy: 82.77%) | | |
| Classified as | Low fat | Med fat | High fat | Low fat | Med fat | High fat | Low fat | Med fat | High fat |
| Low fat | 60 | 0 | 0 | 27 | 11 | 2 | 45 | 0 | 15 |
| Med fat | 0 | 59 | 1 | 6 | 31 | 3 | 3 | 50 | 7 |
| High fat | 0 | 0 | 60 | 0 | 0 | 40 | 13 | 3 | 44 |
| Precision | 1 | 1 | 1 | 0.81 | 0.73 | 0.88 | 0.73 | 0.94 | 0.66 |
| Recall | 1 | 0.93 | 1 | 0.67 | 0.77 | 1 | 0.75 | 0.83 | 0.73 |

We also apply the same procedure with true-color RGB images to compare against the hyperspectral images. We extracted 30x30x3 patches from the im-

ages and calculated the "mean" vector over pixel patches and trained RBF kernel SVMs. Table 3 shows confusion matrices for each food type using RGB images and hyperspectral images. As can be observed from these tables, using hyperspectral images provides us with more discriminant information to classify foods compared to RGB images. However, if we look at the Yams' confusion matrices, it seems that RGB-based features yield higher recall (but lower precision) when testing on the low fat class while hyperspectral features work better for the other two classes. To use the outputs obtained by both set of the features, we used the simple Arithmetic Mean Rule (AMR) [11] to combine the probability outputs of two classifiers. It was found that using AMR, increases the accuracy from 81.66% to 85.0%. The confusion matrix from this approach is given in Table 4. As can be observed from Table 4, in the new approach more samples of low fat and medium fat are classified correctly (yielding higher recall and precision).

**Table 4.** Results for food calorie content identification in yams using RGB and hyperspectral images.

| Classified as | Low fat | Med fat | High fat |
|---|---|---|---|
| Low fat | 29 | 11 | 0 |
| Med fat | 6 | 33 | 1 |
| High fat | 0 | 0 | 40 |
| Precision | 0.87 | 0.75 | 0.97 |
| Recall | 0.72 | 0.82 | 1 |

## 5 Conclusion and Discussion

In this paper we demonstrated that hyperspectral imaging is able to aid in distinguishing between food calorie content with varying percentage of fat. For food type detection, we generated features from the visible spectrum only and achieved a 97% F-measure. When distinguishing calorie content, our method using PCA and SVM with an RBF kernel on hyperspectral images was able to improve our ability to distinguish calorie content, compared to using only true-color RGB images. We further show potential for improving classification by augmenting data from both hyperspectral and RGB images. Such findings show that further expanding existing databases of food images with hyperspectral images may further advance automated image-based calorie detection. However, there remain some limitations in this work, which we are inspired to address in the future. First of all we use a limited number of food types, and only vary fat content. We also only analyze hyperspectral data with wavelengths ranging from 393 to 892nm. Moreover, while our food is cooked in a home environment, our images are extracted in a laboratory environment. Our results inspire further research in the use of hyperspectral imaging to advance automated estimation of calorie content.

# 6   Acknowledgements

# References

1. Marios Anthimopoulos, Joachim Dehais, Peter Diem, and Stavroula Mougiakakou. Segmentation and recognition of multi-food meal images for carbohydrate counting. In *Bioinformatics and Bioengineering (BIBE), 2013 IEEE 13th International Conference on*, pages 1–4. IEEE, 2013.
2. Diwan P Ariana and Renfu Lu. Evaluation of internal defect and surface color of whole pickles using hyperspectral imaging. *Journal of Food Engineering*, 96(4):583–590, 2010.
3. Joachim Dehais, Sergey Shevchik, Peter Diem, and Stavroula G Mougiakakou. Food volume computation for self dietary assessment applications. In *Bioinformatics and Bioengineering (BIBE), 2013 IEEE 13th International Conference on*, pages 1–4. IEEE, 2013.
4. Gamal ElMasry, Ning Wang, and Clément Vigneault. Detecting chilling injury in red delicious apple using hyperspectral imaging and neural networks. *Postharvest Biology and Technology*, 52(1):1–8, 2009.
5. Mohammed Kamruzzaman, Gamal ElMasry, Da-Wen Sun, and Paul Allen. Prediction of some quality attributes of lamb meat using near-infrared hyperspectral imaging and multivariate analysis. *Analytica Chimica Acta*, 714:57–67, 2012.
6. Fanyu Kong and Jindong Tan. Dietcam: Automatic dietary assessment with mobile camera phones. *Pervasive and Mobile Computing*, 8(1):147–163, 2012.
7. Jiangbo Li, Xiuqin Rao, and Yibin Ying. Detection of common defects on oranges using hyperspectral reflectance imaging. *Computers and Electronics in Agriculture*, 78(1):38–48, 2011.
8. Yuji Matsuda, Hajime Hoashi, and Keiji Yanai. Recognition of multiple-food images by detecting candidate regions. In *Multimedia and Expo (ICME), 2012 IEEE International Conference on*, pages 25–30. IEEE, 2012.
9. J Qiao, N Wang, MO Ngadi, A Gunenc, M Monroy, C Gariepy, and SO Prasher. Prediction of drip-loss, ph, and color for pork using a hyperspectral imaging technique. *Meat science*, 76(1):1–8, 2007.
10. Jun Qiao, Michael O Ngadi, Ning Wang, Claude Gariépy, and Shiv O Prasher. Pork quality and marbling level assessment using a hyperspectral imaging system. *Journal of Food Engineering*, 83(1):10–16, 2007.
11. Arun Ross and Anil Jain. Information fusion in biometrics. *Pattern recognition letters*, 24(13):2115–2125, 2003.
12. Geeta Shroff, Asim Smailagic, and Daniel P Siewiorek. Wearable context-aware food recognition for calorie monitoring. In *Wearable Computers, 2008. ISWC 2008. 12th IEEE International Symposium on*, pages 119–120. IEEE, 2008.
13. Fengqing Zhu, Marc Bosch, Insoo Woo, SungYe Kim, Carol J Boushey, David S Ebert, and Edward J Delp. The use of mobile devices in aiding dietary assessment and evaluation. *IEEE journal of selected topics in signal processing*, 4(4):756–766, 2010.