

Abstract

Motivation.

The goal of this project is to produce a method to noninvasively detect allergens in a given food item. The potential market value of such a product is about \$760 million due to the high incidence and cost of food allergens. This project investigates the possible use of hyperspectral imaging, which provides information about the spectral response of food items.

Methods.

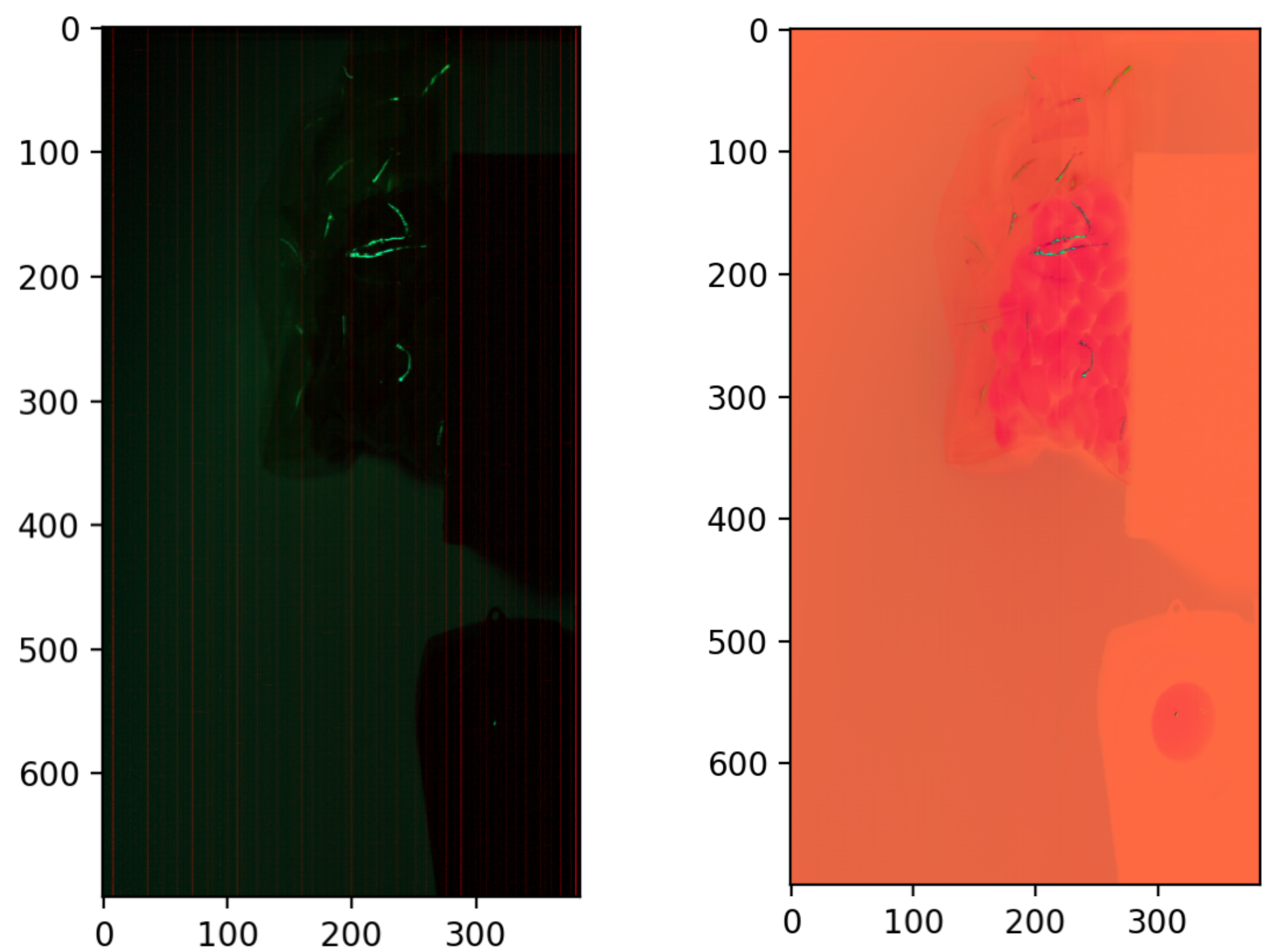
Experiments were conducted using a MCT hyperspectral camera and a FTIR detector to measure the spectra of various food samples across the range of 900-19000 nm. A parameter-based neural network algorithm was used to categorize a given food item and tested for accuracy.

Results.

Hyperspectral imaging in the SWIR range (900-1700 nm) appears to be a suitable solution to identify categories of allergens with a reasonable price. The accuracy of the algorithm is limited in the cases of varying lighting conditions, specific allergen detection, and trace detection. Adjusting the hyperparameters, smoothing the input, and using multi-pixel analysis could increase the accuracy of the algorithm.

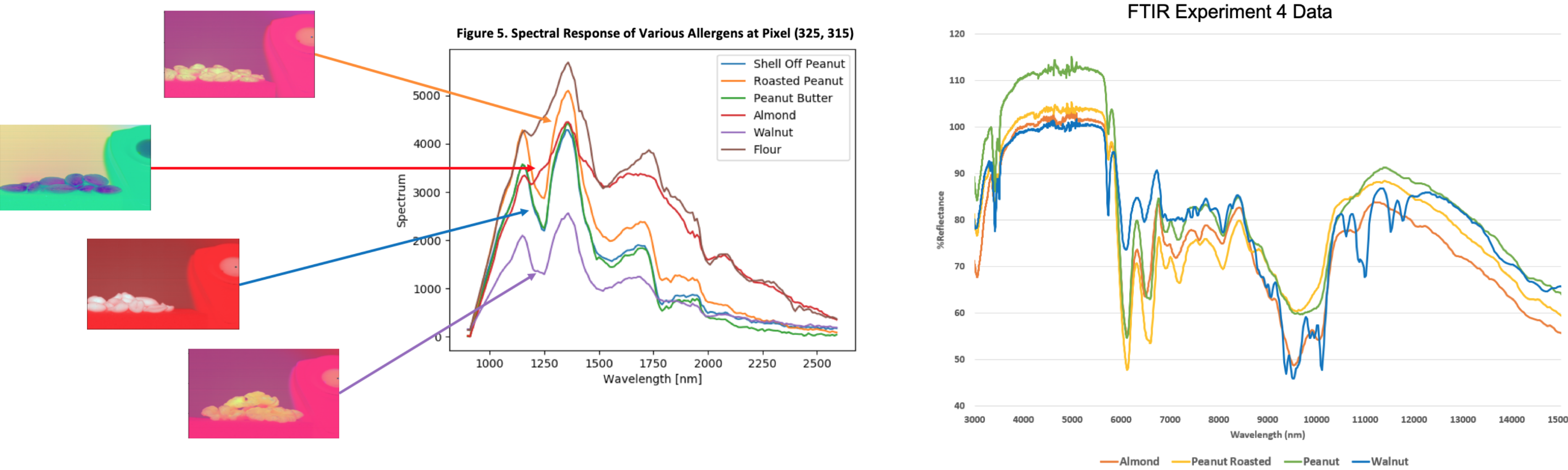
Introduction and Market Research

- 5% of the American population has a food allergy, and the incidence of food allergies has increased by about 50% from 1999 to 2011
- A key allergen of interest is peanuts due to its particularly high prevalence and cost
- Current allergen detection devices in the market are invasive (requires touching or manipulating food sample) and can only detect one particular allergen.
- The projected market value of a device that can non-invasively detect food allergens is estimated to be \$760 million in 2022
- Additional specifications of an ideal product are as follows: fast, portable, able to detect both presence and amount of allergen, and costs on the order of \$100s to \$1000s
- Hyperspectral images (which require a specialized camera) give the spectral response of each pixel across a range of wavelength
- Below is a hyperspectral image of peanuts (left) and the corresponding principal component analysis image (right) for better visualization



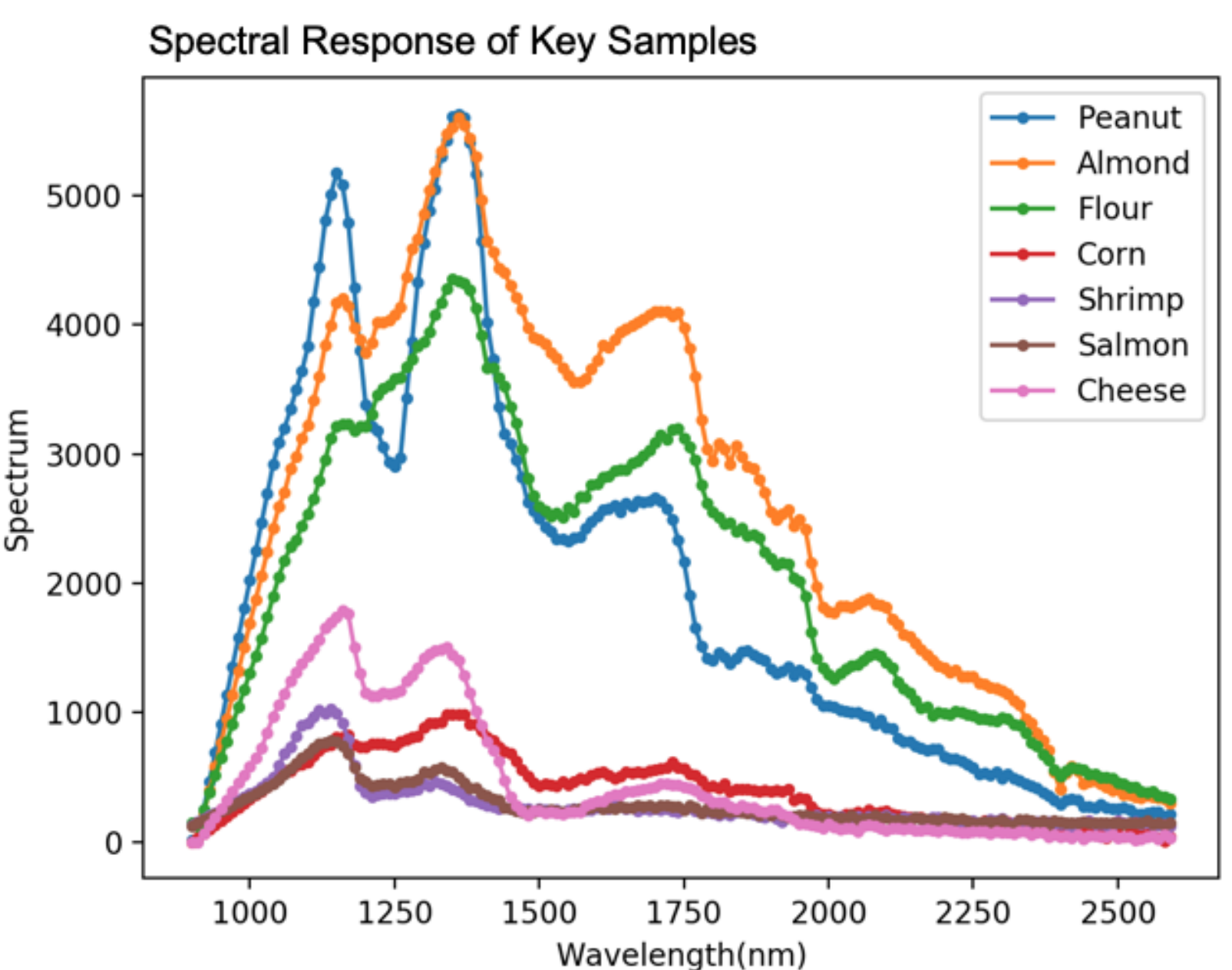
Hyperspectral Imaging and Food Allergens

- Conducted experiments with an MCT hyperspectral camera (Extended SWIR 900-2500 nm, cost on order of \$10000 per detector)
- Limiting the range to SWIR (900-1700 nm) allows for use of a cheaper InGaAs detector (cost on order of \$1000 per detector)
- Also conducted experiments with an FTIR (MWIR/LWIR/FIR 2500-19000 nm) to determine feasibility of a thermal camera/optical filter solution



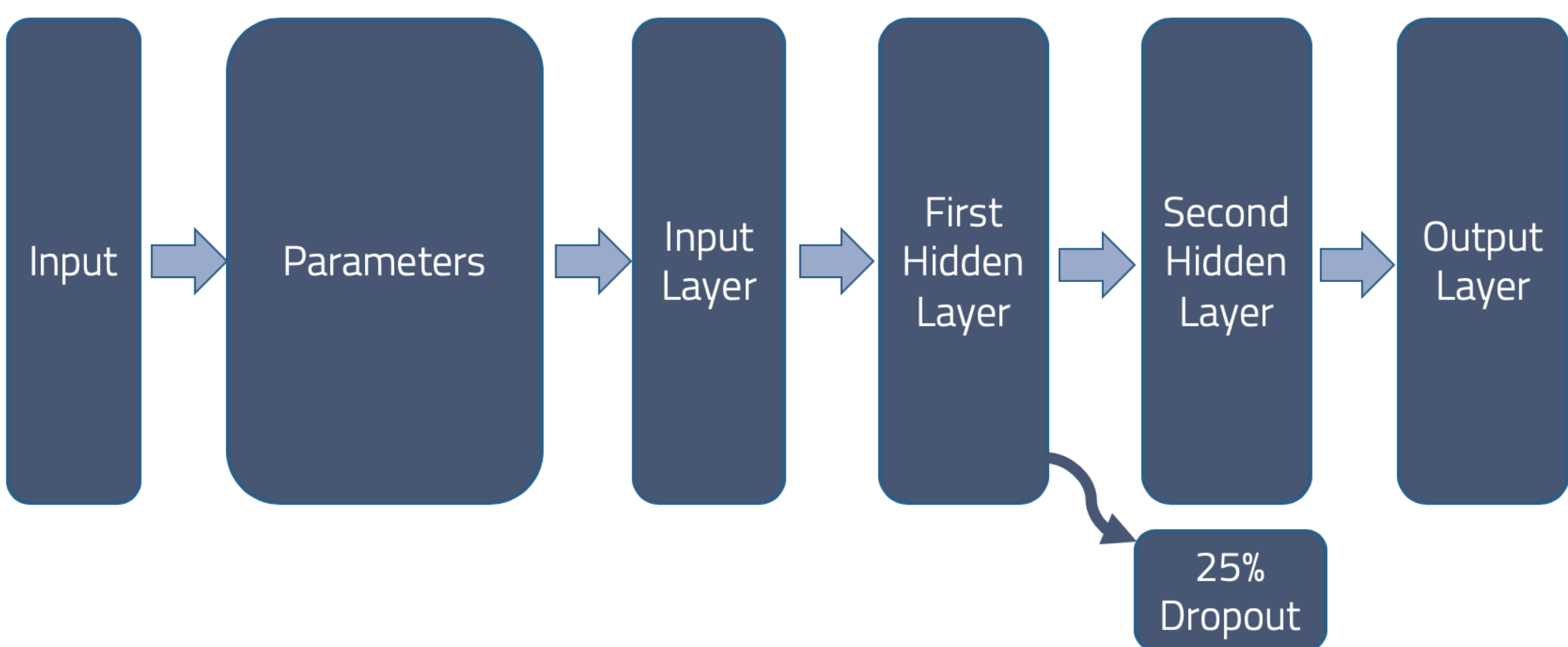
Parameter-Based Algorithm

- Spectrum based input can be inefficient and misleading
- Using parameters is more robust since it reduces input noise and complexity of the neural network
- Parameters are sensitive to lighting conditions and food preparation/cooking but not to temperature
- Calculated as a data preprocessing step then fed as inputs
- Increases accuracy for difficult cases (eg. peanut vs walnut)



Neural Network

- Uses a multi-layer perceptron with 2 hidden layers, ReLU as an activation function and Adam for stochastic optimization.
- Trained with hyperparameters: k-value of 2, batch size of 2, and 100 epochs
- Conducts pixel-by-pixel analysis
- Outputs the probability that a given pixel contains the specified allergen (or for multiple allergens, the probability of each)
- Grouping allergens increases the accuracy, and blurring the input has limited impact



Conclusions

There exists a huge US-centric market for a product that can quickly and conveniently detect allergens in food products. The demand is still high even for a product that can only detect a single allergen (eg. peanuts) or groups of allergens. Hyperspectral imaging appears to be a promising solution. The data suggests that a sufficient amount of distinguishing features are present in the SWIR range (which is within the range of relatively cheaper InGaAs detectors). The prototype parameter-based algorithm (which uses a single architecture for classification of each allergen) has limitations particularly in distinguishing between similar allergens (eg. peanut vs walnut), varying lighting conditions, specific allergen detection, and trace detection. Improving the algorithm adjusting the number of layers and epochs, smoothing the input, and using multi-pixel analysis could increase the accuracy. Trace detection would be necessary for a practical product.

Additional Questions?

Contact the team at ppdd@lists.johnshopkins.edu

Model Description	SWIR Accuracy	Ext. SWIR Accuracy
Peanuts	94.36%	93.31%
Tree nuts	89.60%	99.98%
Nuts	95.16%	96.29%
Wheat	95.20%	88.4%
Shrimp	93.88%	94.06%
Fish	97.88%	98.54%
Dairy	90.20%	94.02%
Egg	92.21%	88.35%
Soy	94.00%	94.45%
Non-allergen	87.83%	89.76%
Everything	74.82%	76.74%
Everything (nut)	73.76%	72.86%
All Allergens	72.19%	83.42%