BabyNutri: A Cost-Effective Baby Food Macronutrients Analyzer Based on Spectral Reconstruction

HAIYAN HU, Hong Kong University of Science and Technology, China
QIANYI HUANG, Sun Yat-sen University, China and Peng Cheng Laboratory, China
QIAN ZHANG*, Hong Kong University of Science and Technology, China

The physical and physiological development of infants and toddlers requires the proper amount of macronutrient intake, making it an essential problem to estimate the macronutrient in baby food. Nevertheless, existing solutions are either too expensive or poor performing, preventing the widespread use of automatic baby nutrient intake logging. To narrow this gap, this paper proposes a cost-effective and portable baby food macronutrient estimation system, BabyNutri. BabyNutri exploits a novel spectral reconstruction algorithm to reconstruct high-dimensional informative spectra from low-dimensional spectra, which are available from low-cost spectrometers. We propose a denoising autoencoder for the reconstruction process, by which BabyNutri can reconstruct a 160-dimensional spectrum from a 5-dimensional spectrum. Since the high-dimensional spectrum is rich in light absorption features of macronutrients, it can achieve more accurate macronutrient estimation. In addition, considering that baby food contains complex ingredients, we also design a CNN nutrition estimation model with good generalization performance over various types of baby food. Our extensive experiments over 88 types of baby food show that the spectral reconstruction error of BabyNutri is only 5.91%, reducing 33% than the state-of-the-art baseline with the same time complexity. In addition, the nutrient estimation performance of BabyNutri not only obviously outperforms state-of-the-art and cost-effective solutions but also is highly correlated with the professional spectrometer, with the correlation coefficients of 0.81, 0.88, 0.82 for protein, fat, and carbohydrate, respectively. However the price of our system is only one percent of the commercial solution. We also validate that BabyNutri is robust regarding various factors, *e.g.*, ambient light, food volume, and even unseen baby food samples.

CCS Concepts: • **Human-centered computing** → **Nutrient estimation**; **Mobile devices**; • **Computing methodologies** → *Machine learning*.

Additional Key Words and Phrases: nutrient estimation, near-infrared spectrometer, baby food, spectral reconstruction

ACM Reference Format:
Haiyan Hu, Qianyi Huang, and Qian Zhang. 2023. BabyNutri: A Cost-Effective Baby Food Macronutrients Analyzer Based on Spectral Reconstruction. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 1, Article 15 (March 2023), 30 pages. https://doi.org/10.1145/3580858

1 INTRODUCTION

Proper nutrient intake is essential for infants and toddlers to ensure children's growth, health, and development to their full potential [24]. Instead, poor diets may contribute to under- and overnutrition, which significantly

*Corresponding author.

Authors' addresses: Haiyan Hu, Hong Kong University of Science and Technology, Hong Kong, China, hhuap@connect.ust.hk; Qianyi Huang, Sun Yat-sen University, Guang Zhou, China and Peng Cheng Laboratory, Shen Zhen, China, huangqy89@mail.sysu.edu.cn; Qian Zhang, Hong Kong University of Science and Technology, Hong Kong, China, qianzh@cse.ust.hk.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.
2474-9567/2023/3-ART15 $15.00
https://doi.org/10.1145/3580858

harm children's health. It is reported by WHO that undernutrition is associated with 45% of all child deaths [38]. Furthermore, overnutrition during infancy could develop into obesity, which increases the risk of severe health conditions, including heart disease, diabetes, and cancer [33]. Therefore, mothers need to know whether the food consumed by children is up to the specified nutrition intake recommendation. Among various nutrition, macronutrients, *i.e.*, carbohydrate, fat, and protein, as the primary energy source, play particularly crucial roles in the baby's growth and development of muscles, bones, and brain tissue [18]. Therefore, tracking and managing the macronutrient intakes of infants and toddlers is especially important.

Currently, mothers can only track their babies' nutrient intake from food nutritional labels (for commercial food) or recipes (for homemade food). Especially for homemade food, mothers need to record the amount of each ingredients and calculate the nutritional intake at every meal, which could be too cumbersome to be sustainable. Although there are some semi-automatic solutions which can help calculate the nutrition, they still ask for user input of ingredients and amount, which is still a tedious procedure for parents [23]. To this end, many automated food nutrition estimation systems have been proposed to simplify this tracking process. One of the mainstream methods of food nutrition recognition is based on computer vision technology [7, 8, 22]. The underlying idea is to leverage the visual features (*e.g.*, shape and color) in food images to infer the food type and search for food composition information from the online database. Nevertheless, baby food is mainly made into pastes, with all ingredients mixed together. Thus it is difficult to identify ingredients in baby food by visual features. Furthermore, there are several optical signal-based solutions that have been proposed to deal with food samples with similar appearances [15, 16, 20, 21, 26]. For instance, Smart-U [15] can achieve accurate food identification by only several light-emitting diodes (LEDs) and photodiodes (PDs). AI-light [21] embeds several LEDs to detect the alcohol concentration of beverages. Although these systems can achieve in-depth analysis of food components, their performance is limited, which can only achieve quantitative or qualitative analysis of a single type of food. Using these systems to recognize the concentration of multiple macronutrients for baby food with complex ingredients is difficult. In addition, albeit some commercial spectrometers can access accurate nutrient measurements [4, 29], these systems often use sophisticated hardware, making them expensive (hundreds of dollars) and difficult to implement in everyday life. As far as we know, there is no automated solution for baby food macronutrient estimation in daily usage. To narrow this gap, we ask whether we can propose a cost-effective macronutrient estimation system for baby food samples that can achieve comparable performance with commercial spectrometers.

We find that the bottleneck limiting the performance of low-cost spectrometers stems from their low-dimensional light spectrum. The low-cost spectrometers reduce the cost by utilizing several LEDs with different wavelengths to substitute the full-band light source. Thus they can directly receive a discrete spectrum without the need to split components that are used to disperse the mixed light. However, this design comes at the expense of dimension reduction of the receiving spectrum. The spectral dimension produced by the system is limited by the number of the light source, *i.e.*, the number of LEDs, which is generally only a few (*e.g.*, 4 [21]). However, because every nutrient tends to have specific absorption characteristics at many different wavelengths, it is not easy to accurately measure all three macronutrients by such low-dimensional spectra. To this end, obtaining a high-dimensional spectrum that contains the fine-grained light absorption characteristics is necessary to estimate macronutrients for baby food.

Fortunately, we observe that each PD can actually detect the light integrated over its spectral-sensitive bandwidth, which typically ranges as wide as more than 100 nm. That is to say, when receiving a mixed light, although the PD only outputs one single value, it actually includes a range of more than 100 nm spectral information. This observation reveals the opportunity to restore a high-dimensional spectrum from the PD's single output. For the same food sample, the difference between the spectra obtained by different systems, *i.e.*, low-cost and commercial systems, is mainly from the parameters between the system hardware. Thus, we can reconstruct the high-dimensional spectra from the PD output by learning a mapping between its output to

the output of a professional spectrometer. This mapping represents the relationship between the two system hardware parameters. Therefore, if we can learn a stable mapping, we can leverage a cost-effective spectrometer to achieve comparable performance with commercial devices on macro-nutrition estimates for baby food.

Based on this observation, this paper proposes a cost-effective baby food macronutrient estimation system, BabyNutri, based on spectral reconstruction. BabyNutri can reconstruct a more informative high-dimensional spectrum from the output of the photodiode array and further deduce the macronutrients of the baby food samples. Different from previous LED-array-based low-cost systems [15, 16, 21], BabyNutri utilizes a full-band light source and a PD-array to obtain the low-dimensional spectra. This PD-array-based design can make the system affordable while taking full advantage of the fact that PD can collect light in a wide band. The obtained low-dimensional spectrum by PD-array can be guaranteed to have the spectral information of the whole band so as to have the opportunity for spectral reconstruction. Nevertheless, to transform our idea into a practical system, we have to address the following challenges:

- *Firstly*, the spectral reconstruction process is non-trivial. Ideally, the reflection spectra of the same food sample collected by two devices are only influenced by the food's reflection rate and the hardware parameters of the devices (*e.g.*, light source energy spectrum, the detector responsivity). So it is easy to learn a stable mapping by eliminating the food reflectance. However, some factors will affect the stability of spectral reconstruction in practice, such as ambient light and surrounding objects scattering. Due to the introduction of these interference factors, the learned mapping is affected not only by the hardware parameters of the two devices but also by the food samples' reflectance and environment. Thus, the learned mapping will overfit the training environments and samples, leading to bad generalization performance.
- *Secondly*, it is challenging to design a nutrient estimation model that can be generalized for all types of baby food. Out of the consideration to supplement multi-dimensional nutrition, baby food ingredients are usually very complex, including fruits, vegetables, fish, meat, grains, *etc* . Most previous models based on the spectrum for predicting nutrition only focus on predicting a single type of food (*e.g.*, corn [10]). So there is no general model for predicting food nutrition with complex ingredients.
- *Thirdly*, collecting low-dimensional spectra that contain full-band spectral information is difficult. As we mentioned, BabyNutri utilizes a full-band light source and an array of photodiodes to obtain the low-dimensional spectra containing the whole band's spectral information. However, two problems still need to be solved using this design. First, the PDs' spectral response ranges are overlapped, so readings from each other may interfere and overlap. Second, the spectral responses of PDs are not of the same magnitude. Thus the readings of some PDs may be too high and dominate the reconstruction process.

We solve the challenges mentioned above in terms of both hardware design and processing algorithms. For the first challenge, we rephrase the problem of learning a stable reconstruction algorithm to the denoising problem of an unstable reconstruction algorithm. To solve this problem, we propose a novel autoencoder-based spectra reconstruction algorithm that combines a linear regression model and a novel convolutional denoising autoencoder model to learn the linear mapping between the spectra while combating the non-linear noise. First, we reconstruct some noise-carrying high-dimensional spectra from the low-dimensional ones using the linear regression model. Then we feed these noise-carrying reconstructed spectra into a denoising autoencoder, which can achieve stable spectral reconstruction by extracting noise-independent compressed representations in the latent space. For the second challenge, we exploit convolution neural networks to extract the features from the reconstructed spectra effectively. To compensate for the nutritional prediction bias caused by reconstruction errors, we pre-train the model on the original high-dimensional spectrum and then fine-tune the model using the reconstructed spectrum. At the same time, we also introduce several data augmentation methods for spectral data, which effectively improve the performance of the nutrition prediction model. Last but not least, we deploy

Fig. 1. Absorption spectra of various macro-nutrients.

a multi-channel ampli er circuit for the PD-array-based device, which can e ectively reduce the interference and dark current noise between di erent branches while amplifying the signals of di erent wavelengths.

We implement a fully-functional BabyNutri that can accurately estimate three macronutrients in various types of baby food. We have conducted extensive experiments on 88 types of baby food with various categories to evaluate the performance of BabyNutri. Experiment results show that BabyNutri outperforms several state-of-the-art baselines on spectral reconstruction with only 5.91% error. Meanwhile, our proposed system's performance on baby food nutrient estimation is obviously outstanding to state-of-the-art cost-e ective solutions. Even when compare with dedicated professional spectrometer, BabyNutri obtain a high correlation with the professional devices, i.e., 0.81, 0.88, and 0.82 for protein, fat, and carbohydrate, respectively. However, the price of our system is only one percent of the commercial one. We also verify that BabyNutri is robust to various environmental factors, such as temperature, illumination, and materials. Furthermore, we explore the ability of BabyNutri to be applied to samples that are not baby food, which indicates the potential of the proposed solution to be further deployed into other areas.

Contributions : We conclude the contributions of this work as follows:

- We design BabyNutri, a low-cost PD-array-based spectrometer, which can deduce baby food nutrient concentration from reconstructed high-dimensional spectra with comparable performance to commercial solutions but signi cantly lower hardware costs.
- We propose a novel autoencoder method to achieve accurate and generalizable spectra reconstruction and a CNN model to deduce nutrient concentration for various types of baby food.
- Extensive experiments over more than 80 types of baby food show that BabyNutri can achieve comparably accurate macronutrient estimates to commercial devices, with correlation coe cients of 0.81, 0.88, and 0.82 for protein, fat, and carbohydrate, respectively. Meanwhile, BabyNutri is adaptable to various scenarios, such as ambient light, materials, and even unseen food samples.

In the rest of the paper, we rst introduce the system's working principle in Section 2. We describe the system structure in Section 3. The detailed design of system is introduced in Section 4, followed by implementation in Section 5 and evaluations in Section 6. We discuss limitations and future works of our system in Section 7 and review the state-of-the-art literature related to our system in Section 8. Section 9 concludes the paper.

Fig. 2. Working principle of di erent spectrometers.

2 BACKGROUND AND RATIONALE

2.1 Nutrients and Near-Infrared Spectroscopy

Near-infrared light refers to electromagnetic waves in the 780-2500nm range, which resonates with chemical bonds, e.g., C-H, H-O, and N-H [25]. These chemical bonds absorb electromagnetic waves of speci c wavelengths, leading to observable intensity absorption bands in the NIR region. Therefore, major nutrients in food, such as water, protein, fat, and carbohydrate, can be identi ed by analyzing the absorption spectrum. Figure 1 shows the near-infrared absorption spectra of foods representing three nutrient elements (whey protein, lard, and glucose) in the 900-1700nm range. We can obviously nd that the three nutrient elements show di erent absorption characteristics in this band. Furthermore, it is also possible to determine the concentration of nutrients of interest by measuring the intensity of re ected or radiated NIR from the food sample. Beer's Law [35] revealed the relationship between the NIR absorption and the concentration of nutrients in a sample, i.e., = Y;2, where and 2 are the absorbance and concentration of the sample, Y is the molar attenuation coe cient and the optical path length of the light, respectively. In conclusion, we can identify and quantify food's nutrient compositions by analyzing its near-infrared spectral absorption.

2.2 Characters of Existing Spectrometer

Most professional spectrometers obtain the food's near-infrared absorption spectra by spectrophotometry, which utilizes a di raction grating to divide the re ected light into several beams with di erent wavelengths. As shown in the left system in Figure 2, a mixed light (containing di erent wavelengths) is emitted from the source and then dispersed into several beams with di erent wavelengths by a di raction grating. The exit slits can pass through a beam of monochromatic light to react with the sample, which is nally detected by the detector. By modulating the pass wavelength of the slits, the detector can eventually gain a complete absorption spectrum.

However, the di raction grating components are cumbersome and expensive, limiting the professional devices only for experimental applications. Some commercial spectrometers simplify the system's structure, which dramatically reduces in size of the equipment [17, 29]. Yet the core of the device still uses the hardware components for optical splitting operation, so the entire device cost is still more than hundreds of dollars. Recent years have witnessed several low-cost spectrometer systems based on LED and PD components [15, 28]. However, as shown in the middle system in Figure 2, these LED-array systems can only obtain low-dimensional spectra, whose dimensions correspond to the number of LEDs, which is generally only a few. Since the low-dimensional spectra only contain limited light absorption characteristics of the nutrients, these systems can not be used for baby food macronutrient estimation. Hence, there is still a critical need for a mobile, a ordable, and accurate macronutrient estimation system for baby food.

2.3 Rationale of Spectra Reconstruction

Figure 2 illustrates the work principles of di erent spectrometer systems. By comparing the working process of professional and PD-array spectrometers, we nd there is an opportunity to achieve a mapping between the two devices which can help the low-cost spectrometer achieve closer performance to the professional one.

Given a speci c food sample (e.g., an apple), it has a constant re ectivity $5^1_^o$. We use two di erent spectroscopy systems to detect the re ection spectrum of the sample. One is a professional spectrometer based-on di raction grating, which can split the light into " sub-channels and introduce a high-dimensional spectrum with " dimensions. Another is a PD-array spectrometer based on # photodiodes, which can generate a low-dimensional spectrum with # dimensions # is much less than " . But the light sources of the two systems cover the same spectral range (e.g.,).

For the PD-array spectrometer, the output of the $_1$ PDs can be depicted as:

$$+_: = \quad _1{}^1_^o 5^1_^o (_: {}^1_^o 3_ = \quad . _: {}^1_^o 5^1_^o 3_ \bullet : = 1 \bullet 2 \bullet \text{"""} \bullet \# \bullet \tag{1}$$

where $_$ is the wavelength, $+_:$ is the output of the PD $:$- in the system, $_1{}^1_^o$ is the spectral power distribution of illumination, $(_: {}^1_^o$ is the sensitivity of the PD $:$-, $5^1_^o$ is the spectral re ection of the sample. $^1_^o$ is the product of the $_1{}^1_^o$ and $(_: {}^1_^o$, which is the spectral responsivity of the PD $:$-The matrix from Eq. 1 can be represented as:

$$+ = . 5 \bullet \tag{2}$$

where $+$ is the vector of PD-array spectrometer output, is the system's spectral responsivity matrix, and 5 is the vector of sample re ection.

For the professional spectrometer, the output of each $_1$ sub-channels is depicted as:

$$+_: {}^0 = \quad _0{}^1_^o) _: {}^1_^o 5^1_^o (_0{}^1_^o 3_ = \quad . _: {}^0{}^1_^o 5^1_^o 3_ \bullet : = 1 \bullet 2 \bullet \text{"""} \bullet \text{"} \bullet \tag{3}$$

where $+_: {}^0$ is the output of the $:$ -th sub-channel in the system, $_0{}^1_^o$ is the spectral power distribution of illumination, $(_0{}^1_^o$ is the sensitivity of the detector), $_: {}^1_^o$ is the spectral transmittance of the lter in subchannel $:$. $^0{}^1_^o$ is the product of the $_0{}^1_^o$, $(_0{}^1_^o$ and $)_: {}^1_^o$, which is the spectral responsivity of each sub-channel. The matrix form of Eq. 3 is then expressed as:

$$+ {}^0 = . {}^0 5 \bullet \tag{4}$$

where $+ {}^0$ is the vector of professional spectrometer output, and ^{0} is the system's spectral responsivity matrix. Thus, to reconstruct the high-dimensional spectrum m^0 (from the low-dimensional one $+$ (), we assume the reconstruction matrix is , . The reconstructed high-dimensional spectrum is expressed as:

$$+ e^0 = , + \text{"} \tag{5}$$

By minimizing the square error between the reconstructed high-dimensional spectra and the professional spectrometer output, we can derive the reconstruction matrix, as shown in Eq. 6:

$$\frac{m k + e^0 \quad + {}^0 k}{m,} = 0 ! \quad , \quad = E^1 + {}^0 +) \, {}^o E^1 + +) \, {}^o \; {}^1 \text{"} \tag{6}$$

As the re ectivity of the food sample is constant, the reconstruction matrix actually indicates the system parameter mapping between the two systems. Thus, the low-dimensional outputs obtained by a PD-array spectrometer can be reconstructed into high-dimensional spectra with this reconstruction matrix. Nevertheless, as there are some dynamic factors, such as scattering and ambient light, between the two systems and the reconstruction and testing phases, it is hard to derive the general reconstruction matrix by Eq. 6. We will analyze how these dynamic factors in uence the reconstruction process and how BabyNutri solves the problems in Section 4.

Fig. 3. The system structure of BabyNutri.

3 SYSTEM OVERVIEW

Previous studies tend to use an array of light-emitting diodes (LEDs) with various central wavelengths as the light source and one or two photodiodes (PDs) covering a wide range of wavelengths as the receiver [15, 16, 21], whose system structure are shown in the middle of Figure 2. Since the LED has a very narrow spectral bandwidth, e.g., the e ective illumination wavelength of a RED LED is from 630 nm to 650 nm, it can be regarded as a monochromatic light source. Turning on the LED array allows the receiver to collect the re ected light at di erent wavelengths, which constitute a low-dimensional spectrum. However, this kind of low-dimensional spectra cannot be used for spectral reconstruction as they only contain information of a few discrete narrow bands in a broad-band spectrum. If we want to obtain a spectrum of 900 nm to 1700 nm, we need about 40 LEDs to cover the full-band spectrum, which is ine cient in both cost and system size.

Therefore, in this paper, we design a novel minimized spectroscopy prototype based on an incandescent lamp and an array of photodiodes. Unlike the LED-array system, our system can produce a full-band spectrum including ultraviolet, visible and infrared like a professional spectrometer. For the receiver, we exploit an array of photodiodes to collect re ected light at di erent wavelength ranges. As photodiodes always have a wide range of spectral responses, such as 300 nm or 400 nm, only a few photodiodes are needed to cover the full-band spectrum. For example, if we want to obtain a spectrum of 900 nm to 1700 nm, we only need about three photodiodes to cover the interested band. In general, we can also appropriately increase the number of PDs to increase the accuracy of reconstruction [12]. As shown in Figure 2, although the PD-array system also generates a low-dimensional spectrum, it actually contains the entire band light absorption characteristics of the food, similar to the professional spectrometer. The di erence is that the light absorption characteristics are compressed into a low-dimensional output. Therefore, it is possible for PD-array system to reconstruct a hundred-dimensional spectrum from this low-dimensional output.

We utilize this spectral reconstruction idea for baby food macronutrient estimation and propose BabyNutri. Figure 3 shows the system structure of the system. The system generates a full-band spectrum and collects the re ected light from the food sample with n photodiodes. The union of the spectral response ranges of PDs covers the whole spectral band of interest (e.g., 900nm 1700nm). The spectral reconstruction module then reconstructs the low-dimensional output from the PD array to the high-dimensional ones. The reconstructed high-dimensional spectra will further be used to estimate the concentration of food's macronutrients, protein, fat, and carbohydrate. We will describe the detailed design of the two modules in Section 4.

Fig. 4. The emi ed light will travel through three light paths to reach the detector.

4 SYSTEM DESIGN

In this section, we will introduce the detailed design of BabyNutri. We will rst elaborate on what factors may in uence the stable spectral reconstruction and how BabyNutri combats the interference (challenge 1). Then, we will describe how to design a nutrient estimation model that can be generalized to all types of baby food (challenge 2). Last, we will introduce the problems we meet when building the low-cost spectrometer and how we deal with them (challenge 3).

4.1 Spectral Reconstruction

Section 2 introduces the rationale of spectral reconstruction in an ideal setup, where the collected spectra are only a ected by the food re ectance (5) and spectrometer hardware parameters (i.e., , (, and) . In this case, we can easily calculate the reconstruction matrix. In other words, when turning on the system light source to scan the food, we expect all the light will be re ected by the food and arrive at the detector (i.e., & = 5, where & is the intensity of the light arrived at the detector. In this way, the high-dimensional spectrum of the sample is:

$$+ {}^{0} = , + \bullet \tag{7}$$

$$({}_{0} \quad {}_{0}) \quad 5 = , \quad (\quad \& = , \quad (\qquad 5 \bullet \tag{8}$$

$$, \quad = ({}_{0} \quad {}_{0}) \qquad {}^{1} \; (\; {}^{1} \bullet \tag{9}$$

where (is the sensitivity of the photodiodes, (,0,) , and (0 are the illumination power distribution, the lter's spectral transmittance, and sensitivity of the detector of the professional device. Since all these parameters are supposed to be constant, the reconstruction matrix , represents a constant transformation from BabyNutri to the professional device. However, this assumption is hard to hold in practice.

Two major unexpected factors will destroy the stable mapping. Initially, in addition to the light generated by the system light source, ambient light will also be mixed into the system (e.g., oodlights, and sunlight. Furthermore, not all the light emitted by the system light source will scan the food and be re ected as expected. Figure 4 shows the longitudinal section of the system, where the container with the food is placed above a black sensing box that contains a light sensing unit. The light generated by the incandescent light source will travel three paths to reach the detector. First, since we point the lamp directly at the food sample, most light will interact with the food samples and re ect back to the detector (i.e., PD-array. Second, few lights may reach the photodiodes directly. Third, some light may be scattered by the surrounding environment, such as the wall of the sensing box.

Fig. 5. The working process of spectral reconstruction module.

Therefore, the actual intensity of the light that reaches the detector is:

$$\&^{0} = U_{1}\ 5, V_{1}, W_{1}\ 5^{0} \tag{10}$$

where $_{1} = \ , \ _{4=E}$ and $U, V, W = 1$. And $_{4=E}$ is the ambient light that mixes in, 5^{0} is the spectral re ectance of the surrounding materials. Thus, the practical reconstruction process could be:

$$+^{0} = , + \bullet \tag{11}$$

$$(_{0} \quad _{0}) \quad 5 = , \quad (\ ^{1}U_{1}\ 5, V_{1}, W_{1}\ 5^{0} \circ" \tag{12}$$

Since both the materials 5^{0}) and the ambient light intensity ($_{4=E}$) are changeable, it is impossible to calculate a constant transformation between the two systems anymore. What's worse, due to introducing these non-linear factors, we cannot eliminate the food re ectance 5 from the reconstruction. Thus, the mapping learned in the training phase includes mapping the two equipment system parameters and introduces the environmental interference term ($_{4=E}$ and 5^{0}) and the re ectance of the food samples 5 in the training set. If the lighting, sensing box materials, and even sample re ectance in the test phase are di erent from those in the training phase, the reconstruction performance will be poor.

We solve this problem in two folds. First, we eliminate the in uence of ambient light by simple physical means. As illustrated in Figure 4, our prototype is an opaque sensing box at the bottom of the food container. Because baby food samples are opaque pastes, most of the ambient light is blocked before arriving at the detector. In addition, we also apply a simple background elimination approach to further enhance the elimination of ambient light. Speci cally, before turning on the lamp, PDs are successively activated to collect a group of background spectra. Then when turning on the lamp, we will subtract the background spectra from the detected food spectra.

Furthermore, to reduce the dynamic interference, we propose a novel spectra reconstruction algorithm based on a denoising autoencoder that can learn the linear mapping between the spectra while combating the non-linear noises. Figure 5 illustrates the structure of our proposed algorithm, which combines a linear regression model and a convolutional denoising autoencoder model. As depicted in Eq. 9, the linear regression can t the mapping between the outputs of the two devices when there is no interference. To this end, we rst input the collected

low-dimensional spectra into a linear regression model to achieve a rough reconstruction. The reconstructed high-dimensional spectra are noted as $\hat{+}_{=>8B43}$. Compared with the real high-dimensional spectra $+_{A40;}$, $+_{=>8B43}$ contain noises from the environment and food samples during the training phase. We can consider $\hat{+}_{=>8B43}$ as the result of over tting the model to the training set. Therefore, we exploit a denoising autoencoder structure to combat this kind of over tting. We take $+_{A40;}$ as the training samples and $\tilde{+}_{=>8B43}$ as the corrupt data of $+_{A40;}$. Then $(+_{A40;}, +_{=>8B43})$ is used as the training sample to estimate the reconstruction distribution of the autoencoder ( ):

$$?_{A42>=BCADZA40;}\left(\hat{+}_{=>8B43} \mid \tilde{+}_{=>8B43}\right) = ?_{A42>=BCADZA40;}\left(+_{A40;} \mid \tilde{+}_{=>8B43}\right) \quad (13)$$

Thus, by minimize the reconstruction loss, we can reduce the dynamic noises that destroy the constant mapping between the training and test phase.

Figure 5 shows the structure of the proposed convolutional denoising autoencoder model. The core of the denoising autoencoder is to learn the compressed representation $\tilde{E}_{=>8B43}$ without noise interference in the latent feature space. Previous studies [39, 40] have proved the e ectiveness of convolutional neural networks in processing spectra data. Thus we utilize a two-layer convolutional structure as the encoder to generate e ective representations. The rst convolutional layer extracts relevant data patterns from noise-carrying reconstructed high-dimensional spectra. The second convolutional layer learns the more complex abstractions of these patterns and generates a compressed representation, taking away noises and unnecessary information. Next, the data will be attened and input to a full connection layer with few neurons to further extract the compressed representation. After that, we restore the spectra from the compressed representation by a transposed convolutional decoder with the symmetrical structure of the encoder. The number of lters in the two convolutional layers is 1 and 20, respectively, which is su cient to enable representation capabilities for the model. Filter size and stride in each layer are optimized according to the input variable numbers and the spectral dimensions, which are (20,10) in the rst layer and (3,2) in the second layer. Additionally, we nd that using ReLU as the activation function will make some bands be assigned zero in the compressed representation, which makes it di cult for the decoder to recover the original spectrum. So we use LeakyRelu to avoid information loss during coding.

After reconstructing the high-dimensional spectra, we nd that although the reconstruction performance is not bad, it is poor to estimate the food nutrients. By in-depth analysis, we nd the reason is that we take the mean square error (MSE) as the objective function for the autoencoder model, ignoring the shape consistency between the two spectra. Thus, we propose to use the weighted sum of MSE and a shape consistency loss. Speci cally, the shape consistency loss is de ned as the variance of the di erence between the reconstructed spectra and real spectra. Thereby, the loss function for the convolutional denoising autoencoder model is:

$$!_F = \backslash \frac{1}{"} \tilde{O} \; k\hat{E}_8 \; E_8 k^2 \; ¸ \; ¹ 1 \; \backslash \; º \frac{1}{"} \tilde{O} \; k \; \hat{E}_8 \; \; \hat{E} k^2 • \quad (14)$$

where $\backslash$ is the weight of the MSE loss item, and $\hat{E}_8$ and E_8 are the 8-th dimension of the predicted high-dimensional and real high-dimensional spectra, respectively. $\hat{E}_8 = \hat{E}_8 \; E_8$ and $\hat{E} = \frac{1}{"} \; ¹ \hat{E}_8 \; E_8 º$. By using this modi ed loss function, the auto-encoder model can reduce the distance between the reconstructed spectrum and the real spectrum and ensure that the shapes between them are as consistent as possible so that the reconstructed spectrum can be better used for subsequent nutrient composition prediction.

4.2 Nutrient Estimation

Previous studies have validated the e ectiveness of predicting food nutrients by deep learning models [43]. The deep learning approach requires no complex preprocessing of input spectra, saving labor and time and improving model performances. Various nutrient estimation models have been proposed, but most focus on a single type of food with similar nutrient levels, such as corn [10], or meat [42]. However, the baby food samples have various ingredients, such as fruits, vegetables, meats, and creamy, leading to di erent levels of nutrients. Thus, previous

Fig. 6. The architecture of nutrient estimation model.

models cannot be directly applied to our baby food nutrient estimation problem. To this end, we design a novel CNN-based model for baby food nutrient estimation.

We refer to some previous design suggestions for the CNN-based nutrition analysis model, simplifying the number of model layers and increasing the convolutional kernel size [46]. At the same time, Considering the diversity of the nutrients in baby food is signi cant, the main target of our proposed nutrient estimation model is to avoid over tting. Figure 6 shows the architecture of our proposed nutrient estimation model. The model adopts two convolutional layers, a global average pooling (GAP) layer, a batch normalization (BN) layer, and an output layer. The rst convolutional layer smooths the spectral data and extracts informative features. The second convolutional layer learns the more complex abstractions that can best explain the input food spectra. Then the model leverages a GAP layer to replace the fully connected layer to map the extracted features to the baby food nutrients value. Using the GAP layer, the model can reduce the number of parameters, thus preventing over tting. Subsequently, we apply a BN layer after the GAP layer to normalize the output. Using the BN layer causes all samples in a mini-batch to be associated together so that the output of each sample is a ected by all the data in the batch, which is equivalent to an indirect data enhancement to prevent over tting.

By carefully designing the model structure, our proposed nutrient estimation model can learn good results even when the gap in nutrient concentration in training data is signi cant. However, the test set's generalization performance is still not ideal. Considering that there are also some physical and environmental factors resulting in spectral di erences, such as sample particle size, temperature, etc., we further adopt data augmentation on the training spectra to improve the robustness and generalization performance of the network. The core idea is to expand the number of training samples by adding various types of noise to simulate various expected disturbances. For example, we can simulate temperature variation by shifting the peaks to higher or lower wavelengths [27]. In line with this idea, we select four types of noise most likely to appear in our system for augmentation. According to the introduction in [27], we de ne the four types of noise to raw spectra (as follows:

Additive noise: $+_{>5\,5} = + \;\; X_0$, which simulates the baseline shifts or background di erences of the spectra.
Multiplication noise: $+_{<;?} = + \;\; X_1$, which simulates the scatter noises causing by the various optical path lengths or particle sizes.
Shift noise: $r +_{B\,5} = r + \;\; X_2$, which simulates the shift position and intensity of the near-infrared absorption bands due to the temperature changes.
Gaussian noise: $+_{6DB} = + \;\; N^{1}\,`\bullet f^{2o}$, which simulates the random noises causing by device noise.

Applying the above augmentation to the raw training spectra set not only increases the number of the training samples but also improves the robustness of the nutrient estimation model. Therefore, the pre-trained nutrition estimation model has a higher tolerance for the reconstruction error existing with the reconstructed spectrum, so

Fig. 7. Components of the spectral data augmentation.

it can converge faster and achieve comparable performance to the results predicted on real high-dimensional spectra.

Furthermore, when training the model, we nd that the model trained only on reconstructed spectra tends to perform poorly on the test dataset. This may be because the model training process relies too much on the error information in the reconstructed spectrum. Hence, the training set of the reconstructed spectrum is biased in the wrong direction. Therefore, to ensure the model learns in the right direction, we rst train the model on spectra collected by the professional spectrometer. Then the pre-trained model will be ne-tuned on the reconstruction spectra dataset. The rst layer of the pre-trained CNN model will be frozen to ensure the model pays attention to the correct wavelength band.

4.3 PD-Array Spectrometer

The low-dimensional spectrum used to reconstruct is collected by a PD-array spectrometer. However, it is non-trivial to design the circuit of such a spectrometer. Ideally, as illustrated in Figure 8(a), the light source is always on, and the PDs are connected to the output amplifying circuit, in turn, to receive the re ected light and produce an output. However, two problems are facing us when practically designing the prototype circuit.

First, it is hard to physically isolate the PDs that are not activated from the output circuit. Suppose the inactive PD is directly put in a suspension state (continuously input level). Since the spectral response ranges of PDs overlap, the inactive PD can also output currents. In that case, the branch may still have a weak current, making the actual output containing the current from undesirable PDs. To solve this problem, we design a PD-isolation circuit using a PNP triode and a P-Channel MOSFET, shown in Figure 8(b). Among them, the MOS tube controls whether PD is in the activated state, and the triode controls whether the P-Channel MOSFET is in the saturation region. Compared with the direct use of triode, this design can e ectively avoid the suspension state of PD when given low voltage and also ensure that PD is in the bias mode when activated. Under the bias mode, PD has higher sensitivity and can acquire richer food light absorption characteristics.

Another challenge stems from the unbalanced low-dimensional spectra. For instance, the output of PD 8 could be 4.5 V, while the output of PD 9 is only 0.005 V. This is because of the illumination's uneven power distribution and the di erent sensitivity levels of PDs, i.e., responsivity and active area. To solve this problem, we amplify each PD to a similar output level and then connect it to the output ampli er circuit. For instance, the resistor R_4 in Figure 8(b) is the particular amplify resistor for PD 1. Figure 8(c) shows the output amplify circuit. Thus, for

(a) Ideal Circuit (b) PD-Isolation Branch (c) Output Amplify Circuit

Fig. 8. Design of the prototype circuit.

Fig. 9. Prototype of BabyNutri. Fig. 10. Optical sensing unit. Fig. 11. Spectral responsivity of the photodiodes.

the current $_3$ generated from % 1, the nal output could be:

$$+_{>DC} = \frac{3}{1, \frac{'_7}{'_4}} \; '_7 \; {}^{1}1, \frac{'_5}{'_6} o" \tag{15}$$

Thus, for each PD, we can determine the value of $'_4$ to amplify them to the same level and further amplify the output to make them easily detected. Furthermore, to ensure that $'_4$ does not in uence the ampli cation of other branches, we add a diode ($_1$, 1) for each branch to isolate the possible recoil current generated by other PDs.

5 IMPLEMENTATION

We have designed and implemented a compact and low-cost (less than 10 dollars except for the control unit) prototype using o -the-shelf hardware components. The prototype comprises a sensing box, an optical sensing unit, and a computation unit.

3D printing sensing box. We design a 3D sensing box without the lid using the online tool Tinkercad. The model is printed using black Polylactic Acid (PLA). The box size is 47.5mm 49 mm 15 mm, which is suitable for placing the optical sensing unit. As shown in Figure 9, the paste baby food sample is placed in a sealed disposable transparent plastic bowl on the sensing box. The optical sensing unit's light shines through the bowl's bottom to scan the food. The diameter of the bottom of the bowl is diagonally equal to the box so that no ambient light can penetrate the gap into the sensing unit. It should be noted that the sensing box and container can be set up as a whole for practical use, as long as the light emitted by the sensing unit is not blocked.

Optical sensing unit. We implement the emitter circuit and the detection circuit on a 4 cm x 4 cm two-layer print circuit board, shown in Figure 10. We choose an incandescent lamp (VCC7216-ND) as the light source to generate a full spectral illumination spectrum. The power distribution of the incandescent lamp can cover the UV, visible, and infrared bands. To prevent the pins from shaking, we use hot glue to hold them at a high temperature to create a stable light source. Furthermore, we deploy six commercial photodiodes with di erent spectral ranges uniformly surrounding the light source. The PDs are organized in a circle, and each PD is separately connected to the output ampli er circuit. The PDs are distributed evenly around the light source to guarantee that the light goes through equivalent paths before reaching the detector. Considering that most of the o -the-shelf spectrometers [17, 29] estimate the food contents based on the absorption spectra ranging from 900 nm to 1700 nm, we target to reconstruct the high-dimensional spectra of this range. To this end, we select ve commercial PDs whose spectral ranges are 400 nm 1100 nm (209-040-101-411CT-ND), 700 nm 1070 nm (732-13405-1-ND), 700 nm 1100 nm (3147-B1501PD H9B000113U1930CT-ND), 800 nm 1700 nm (209-0090-3111-185CT-ND), 800 nm 1700 nm (209-0090-3111-185CT-ND), respectively. Figure 11 shows the spectral responsivity of the PDs. We also deploy a photodiode (PD1) covering the visible range, i.e., 400 nm 700 nm (751-1051-1-ND), to sense whether there is food in the bowl. Since most ambient light sources cover the visible band, such as uorescent lamps, we can judge whether there is food blocking the ambient light by the readings of the PD1.

The light emitted by the incandescent lamp reaches the food and experiences di use re ection. The photodiode detects the intensity of re ected light and converts it into electrical currents. To reduce the scatter noise produced by dark current and eliminate low-frequency noise, we set the photodiode zero biased, operating in photovoltaic mode. As introduced in Sec. 4, we design a PD-isolation branch circuit for each PD to prevent interference with each other. The core components are a PNP transistor (SS8050-G) and a P-MOS transistor (SI2300DS-T1-GE3) that isolate the current in the branch when the PD is not activated. Subsequently, the electrical current passes through an output amplifying circuit (AD8505ARJZ-R7CT-ND) and is ampli ed to generate a voltage measured as the intensity signal.

Computation unit. We use an Arduino Mega to control lamp illuminating and PD activation and digitize analog signals from the photodiodes. We alternately give each photodiode a high (active) of 100ms and a low (inactive) of 100ms to collect re ected light and background noise, respectively. The incandescent lamp is rated at 5 , and we use the 5 port on the controller to power it.

6 PERFORMANCE EVALUATION

This section evaluates both the spectral reconstruction and nutrient estimation performance of BabyNutri. We evaluate the performance of BabyNutri by answering four key questions: 1) Does BabyNutri achieve better performance of spectral reconstruction than baselines? (Section 6.2.1) 2) Can BabyNutri achieve robust spectral reconstruction under unseen food samples (Section 6.2.2) and various environments (Section 6.2.3)? 3) Can BabyNutri achieve good macronutrient estimation performance using the reconstructed spectra (Section 6.3.1)? 4) Can BabyNutri generalize to unseen food samples for nutrient estimation (Section 6.3.3)?

6.1 Study Setup

6.1.1 Data Collection. We collect 88 commercial baby food samples from famous baby food brands, including 69 instant purees and porridge food samples, ve rice our samples, and 14 diluted samples. All three baby food stages are considered, i.e., purees (4 to 6 months), thicker consistency (6 to 9 months), and soft chewable chunks (10 to 12 months). The sample may contain only a single ingredient, e.g., (mashed peas) or multiple ingredients (e.g., apple and pea puree). We roughly classify the samples into ve classes based on their main ingredients: fruit, vegetable, meat, cereal, and creamy. Table 1 shows the number of the ve food classes for each brand. The baby food samples are put in a sealed disposable transparent plastic bowl and scanned by our proposed prototype. We

Table 1. Statistics on the number of di erent brands and classes of food.

Brand	Number of di erent food classes
HEINZ (#1)	Fruit (6), meat (5), vegetable (3)
LITTLE FREDDIE (#2)	Creamy (4), fruit (6), meat (3), cereal (3), vegetable (7)
EARTH'S BEST (#3)	Fruit (3), meat (2), cereal (2), vegetable (5)
ELLA'S KITCHEN (#4)	Creamy (3), fruit (3), meat (2), vegetable (10)
BABY BASIC (#5)	Cereal (11)
PIGEON (#6)	Cereal (10)

(a) P-F (b) P-C (c) F-C

Fig. 12. Nutrients distribution of the baby food samples.

scan each sample ve times, generating a 5-dimensional low spectrum each time. The experiments are conducted in a room with a temperature of 25 C 28 C. The lighting condition is in line with normal o ce conditions, whose light intensity is about 100 lx 250 lx.

6.1.2 Ground Truth. The ground truth of nutrient values is collected from the labels on the packages. Speci cally, for instant food, we directly use the nutrient facts labeled on the packages as the ground truth. For several rice our food that need to be brewed, we brew them according to the 1:3 mass ratio of food and water and calculate the ground truth nutrient values based on the ratio. In addition, to enrich the dataset scale, we select 14 samples diluted with water for a 1:1 mass ratio, and the nutrient concentration of the diluted sample is recorded as 1/2 of the original sample. The range and distribution of the three nutrients are shown in Figure 12. It can be seen that, di erent from previous studies [10, 42], baby food has a wide range of three nutrients, especially carbohydrates, which ranges from $2.25 \cdot 10^{-6}$ to $17.17 \cdot 10^{-6}$. Furthermore, for each sample, we collected the ground truth high-dimensional spectra by a PerkinElmer LAMBDA 950 UV/Vis/NIR Spectrometer. We use the integrating sphere accessory to collect the di use re ectance spectra generated from the bottom of the sealed samples. The spectra range from 900 nm to 1700 nm with a resolution of 5 nm.

6.1.3 Metrics. To intuitively demonstrate BabyNutri's spectral reconstruction performance, we use the relative error (RE) as the metric. The relative error for each sample is calculated as:

$$A4;0C8E4\ 4A=A>A\ \frac{\sum_{=1}^{"} |E_{A42>=} - E_{A40;}|}{\sum_{=1}^{"} E_{A40;}} \times 100\% \tag{16}$$

where " is the dimensions of the high-dimensional spectra, $E_{A42>=}$ and $E_{A40;}$ are the sample's reconstructed spectrum and ground truth spectrum, respectively.

For the nutrient estimation, we exploit the root-mean-square error (RMSE) as the metric to demonstrate the models' performance. Since the RMSE can not be used to compare data with di erent scales (e.g., the RMSE of carbohydrates will be larger since its ground truth values are higher), we introduce the Pearson correlation coe cient (R) to achieve the comparison of the three nutrients.

6.1.4 Baselines. We take typical spectral reconstruction algorithms Wiener Estimation (WE) [12], Adaptive Wiener Estimation (AWE) [30], Linear Regression (LR), and Weighted Linear Regression (WLR) [11] as the baselines for spectral reconstruction. As for nutrient estimation, we take both high-dimensional spectral baselines and low-dimensional spectral baselines to evaluate the e ectiveness of BabyNutri. For high-dimensional spectral baselines, we substitute our proposed modules with other state-of-the-art algorithms to show the e ectiveness of each proposed module. For low-dimensional spectral baselines, we simulate state-of-the-art LED-array systems and powerful o -the-shelf ICs by selecting speci c channels from our ground truth spectra.

- LRDAE, MLDNN: We use our proposed spectral reconstruction model (LRDAE) combined with a state-of-the-art nutrient estimation model (MLDNN [1]) as the rst baseline. The MLDNN model is implemented according to the parameters described in the article [1]. Since we want to output prediction values for each food sample, we delete its error avoidance module.
- LR, CNN: We use a typical spectral reconstruction model (LR) combined with our proposed CNN-based nutrient estimation model as another baseline.
- GT Spectra, CNN: We use the ground truth high-dimensional spectra directly input to our CNN-based nutrient estimation model as the third baseline. Note that although the performance of BabyNutri is supposed to be inferior to this baseline, BabyNutri far outperforms this solution in terms of price and portability.
- Smart-U [15]: We select six wavelengths used in [15] from our ground truth spectra, namely 860nm, 940nm, 1050nm, 1200nm, 1300nm, 1450nm, 1500nm and 1650nm. The spectra are then processed by principal component analysis (PCA) and seven di erent regression models (i.e., LinearRegression, PLSRegression, K-Nearest Neighbor, Support-Vector-Machine, Random-Forest, XGBoost, and Decision-Tree) to predict the macronutrient concentration. We nally select the outputs from the best regression model as the nal results.
- Lili [16]: We select six wavelengths used in [16] from our ground truth spectra, namely 940nm, 1070nm, 1200nm, 1370nm, 1300nm, 1460nm, 1550nm, 1600nm, 1700nm. The spectra are then processed by PCA and seven di erent regression models, and nally output the best of the seven regression model as the results.
- AS7421 [2]: Since the spectral range of AS7421 is 750nm 1100nm, and the spectral range of BabyNutri is 900 1700nm, we choose 30 wavelengths within 900 1100nm intersecting between the two as the baseline. The spectra are then processed by PCA and seven di erent regression models, and nally output the best results of the seven regression model.
- Smart-U+Lili: We get this enhanced baseline by combining the channels of the Smart-U and Lili. The combined system has 13 channels ranging from 900nm 1700nm.
- Smart-U+Lili+AS7421: We get this enhanced baseline by combining the channels of the Smart-U, Lili, and AS7421. The combined system has 41 channels ranging from 900nm 1700nm.

(a) 5-fold cross-validation performance comparison (b) Average performance and processing time comparison

Fig. 13. BabyNutri spectra reconstruction performance of 5-fold cross-validation comparison with baselines.

BabyNutri (5-channel PD array): We take BabyNutri without spectral reconstruction as another baseline. The 5-dimensional outputs of PD-array are directly input to the PCA and regression models. We nally select the best results of the seven regression model as the baseline.

6.2 Spectra Reconstruction Performance

6.2.1 K-Fold Cross Validation. Since BabyNutri needs high-dimensional spectra collected by professional equipment to assist training, BabyNutri needs to have good generalization to samples that are not considered during the training phase. To evaluate this capability, we begin with examining the overall performance of BabyNutri by adopting 5-fold cross-validation. We divide the samples into ve roughly equal subsets. We use 4 of them each time to train the model and the rest for testing, then rotate the datasets. Figure 13(a) shows the relative error of our proposed method and four baselines. The average RE of our proposed spectral reconstruction model is only 5.91%, which outstands all of the baselines. We note that both the Wiener Estimation and Adaptive Wiener Estimation show poor reconstruction accuracy. Its poor performance stems from the signi cant di erences between the spectra of the baby food sample set. Thus it is not only hard to calculate the wiener matrix in the training phase but also hard to adapt the trained wiener matrix to test samples. The weighted linear regression model improves the performance on the test set to some extent by weighting the samples that are similar to the test samples in the training set. Nevertheless, WLR takes too much time to weigh the sample set for each test sample, which prolongs the execution of the system and hinders practical application. Instead of weighing the sample set, LRDAE combines the linear regression with a denoising autoencoder model that can adapt to various sample spectra while maintaining good reconstruction e ciency. We summarize the average processing time of each method for every fold in Figure 13(b). It can be seen comparing with WLR, LRDAE only takes one-third of reconstruction time, i.e., 0.87s vs. 2.98s, while achieving comparable or even better reconstruction accuracy than WLR.

6.2.2 Generalize to Various Food Features. As there are various types of baby food, it is essential for the reconstruction model of BabyNutri to have a good generalization ability for di erent food samples. Thus, we evaluate the performance of the spectral reconstruction model over baby food with diverse features, including color, viscosity, type, and brand.

(a) Color

(b) Viscosity

(c) Type

(d) Brand

Fig. 14. Spectral reconstruction performance on generalization to various food features.

Food Color. Color may a ect the light re ection characteristics of food, e.g., red food re ects more red light. To evaluate the model's capability to generalize food with di erent colors, we divide 88 food samples into ve categories: yellow, white, green, red (pink/purple), and orange, including 24 yellow, 23 white, 9 green, 5 red, and 27 orange. Food samples of the same color are used as the test set, and other samples are used as the training set. Figure 14(a) shows the spectral reconstruction error on various food colors. The results indicate that our proposed LRDAE model can generalize well to various food colors, in line with our intuition. This may be because we ignore the visible band during the reconstruction process. Various food colors will not in uence the reconstruction on 900 nm to 1700 nm.

Food Viscosity. We further evaluate the reconstruction performance on a di erent level of viscosity. Babies of di erent ages need to consume baby food with various viscosity levels, so BabyNutri need to be capable of adapting to food with di erent viscosity levels. We classify the food samples into three viscosity levels according to the baby food stages, i.e., purees (4 to 6 months), thicker consistency (6 to 9 months), and soft chewable chunks (10 to 12 months). Stage 1 contains 27 food samples, including 14 samples diluted with water. Stage 2 and stage 3 have 48 and 13 samples, respectively. Food samples of the same viscosity level are used as the test set, and other samples are used as the training set. The spectral reconstruction performance in Figure 14(b) indicates that the LRDAE model shows di erent performance on food with di erent viscosity levels. In speci c, the higher viscosity of the food, the better the spectral reconstruction performance. However, when we look into the reasons, we nd the poor performance on stage 1 food samples mainly because of those diluted samples. Since the water and food