Predicting Macronutrient of Baby Food using Near-infrared Spectroscopy and Deep Learning Approach

M N Aulia\textsuperscript{1}, M L Khodra\textsuperscript{2}, A P Koesoema\textsuperscript{3}
Institut Teknologi Bandung, Jalan Ganesha, Bandung, Indonesia
Email: \textsuperscript{1}masyithahaulia@students.itb.ac.id, \textsuperscript{2}masayu@stei.itb.ac.id, \textsuperscript{3}allya.paramita@stei.itb.ac.id

Abstract. —In Indonesia, malnutrition and overnutrition can still be found especially among children and toddlers who are supposed to get adequate nutrition. The cause of malnutrition is inadequate nutrition intake, both in terms of the quantity and quality of the food. Therefore, mothers need to know whether the food consumed by children has reached the specified nutrition intake recommendation. This research develops a system to predict macronutrient content in baby food using Near-infrared Spectroscopy (NIRS) to obtain the spectral profile of the food and Deep Learning approach such as Deep-belief network (DBN) and Convolutional Neural Network (CNN) to build the prediction model. We used instant-porridge scan data from SCiO to build the model. CNN managed to give the best performance with error for carbohydrate, protein and fat 11.70\%, 26.14\% and 28.72\% respectively.

1. Introduction

Food intake is one of the keys to have healthy life [1]. Food intake will greatly affect human health conditions especially among children and toddlers who are still growing. But in reality, malnutrition and overnutrition can still be found especially among children and toddlers who are supposed to get adequate nutrition. The cause of malnutrition is inadequate nutrition intake, both in terms of the quantity and quality of the food. According to WHO, around 45\% children under 5 in the world, suffer and die from malnutrition [2]. In Indonesia, malnutrition can still be easily found. In January 2018, dozens of children in Asmat regency die from malnutrition and hundreds of them need intensive care [3].

On the other hand, it is difficult to determine whether the food consumed by the children has reached the specified nutrition intake recommendation. The ministry of health in Indonesia has issued a nutrition adequacy rate which provides the daily minimum nutrition intake for both children and adults. However, this method is complicated and time consuming as the mothers need to remember the food intake and manually calculate the nutrition intake from the food. Moreover, another concern is the quality of the food consumed by children cannot be calculated using this approach.

Predicting nutrition intake from food using technology has become a hot issue these days in food industry [4]. These researches include camera-based techniques and sensor-based techniques [1]. One of the most popular techniques is NIRS due to its ability to analyze chemical compounds in a food within a short time and its non-destructive way [5]. NIRS analyzes the spectrum of absorption or reflectance of electromagnetic waves with wavelength ranging from 700 to 2500 nm where most of
chemical and biochemical compounds have unique absorption bands so that it can be used to detect macronutrient such as carbohydrate, protein and fat [1].

![Figure 1. The SCiO Scanner](image)

One of the portable NIR scanner to be easily used in daily needs is SCiO due to its handy size. Although SCiO has narrower wavelength range, but it is still suitable for daily needs [1]. SCiO can only analyze wavelength ranging from 740 to 1040 nm. They also provide mobile application so people can use SCiO to scan their food and get information about nutrition in the food using their application.

SCiO works by scanning the food to get the reflected spectrum. People can place the sample of food in a steady surface and hold the scanner about 5-10 mm above the food. If the food is too small and dry, they can place the sample within the solid sample holder for more accurate scan. SCiO will start analyzing the food when the button on the SCiO or scan button in the mobile application is pressed. The reflected spectrum can be found in the web application. These reflected spectra are different for various foods due to different absorption levels [1].

In this paper, we use SCiO as the scanner to build a macronutrient prediction system for baby food. We aim to develop the model using CNN and DBN. We use instant-porridge as the data. The instant-porridge is cooked firstly and then we use SCiO to scan the food. We use 30 instant-porridges in the market as they already provide nutrition facts printed on the package. These nutrition facts are the labels for our predicting system. We only use carbohydrate, protein and fat from the nutrition facts to train our prediction model and correlate the spectra features to their corresponding carbohydrate, protein and fat contents. Finally, we evaluate this prediction system error using RMSE and error.

2. Related Works
There are already some researches in the past that utilize NIRS and machine learning to both classify and regress.

One of the researches use SCiO to predict energy and carbohydrate content from drinks. Thong dkk. (2017) implement SVR and PLS then compared the result [1]. SVR with RBF Kernel achieves the best result compared to PLS.Khan dkk. (2018) use NIRS to predict wind power using DBN [8]. Yang dkk. (2017) also use DBN to detect counterfeit drugs [12]. Li dkk. (2017) also use DBN to predict moisture, protein and ash content from wheat [13]. Lengkey dkk. (2013) implement regression for Jarak Pagar beans to detect the level of water and fat using PLS [14]. There is also a research that implements CNN conducted by Acquarelli dkk. (2017) to classify various foods and drinks dataset [15]. Besides Thong dkk. (2017), other researches use different NIR Scanners, usually NIR scanners developed to be used by laboratories and factories which also means have wider wavelength spectrum. From those previous works, we conclude that we can build a prediction system using DBN and CNN to predict the carbohydrate, protein and fat content from food, in this case, baby food.
3. System Design

3.1. Dataset Preparation
The dataset consists of 450 instances from 30 instant-porridges. An instance is a scan from a bowl of porridge. We also observed different levels of consistency by pouring more or few waters to the porridge. We use instant-porridges for children under 3 and mostly for toddlers above 6 months as they already need complementary foods beside consuming breastmilk.

Macronutrient value from nutrition facts on the package are transformed into the percentages of its content divided by the total weight of the porridge (the weight of water and porridge). The percentages are used as labels for training phase. Table I shows the example of three instances of data from SCiO. The dimension of raw data is 330.

| Spektrum 740   | Spektrum 741   | Spektrum 742   | Spektrum 743   | Spektrum 744   | Spektrum 745   | Spektrum 746   | Spektrum 747   | Spektrum 748   |
|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| 0.8533703     | 0.85370379    | 0.85422415    | 0.71406316    | 0.7139301     | 0.71283065    |               |               |               |
| 0.7645731     | 0.76521258    | 0.76593314    | 0.63705830    | 0.63694547    | 0.63598849    |               |               |               |
| 0.8722648     | 0.87259958    | 0.87316037    | 0.73208251    | 0.73193435    | 0.73081943    |               |               |               |

3.2. Regression models
We use CNN and DBN as some of deep learning methods to achieve better results compared to conventional methods such as Linear Regression and PLS. By using deep learning approach, the model can learn and extract more complex information and features from the data [6].

DBN is one of semi-supervised learning techniques as it combines supervised and unsupervised learning. DBN consists of two main processes which are pre-training using RBM and fine-tuning using Backpropagation. In pre-training phase, random weights and biases will be trained to construct the input data [8]. This phase does not need any labels so it called unsupervised technique [7]. After some iterations, the trained weights and biases will be used in backpropagation and fine-tuned using labels.

CNN is a variation of Multilayer Perceptron. There are two additional layers in CNN namely convolutional layer and pooling layer. These layers extracts feature from the data to make more representative feature that can help model to learn and build better model [9]. The feature then will be trained in fully-connected layer which consists one or more hidden layers. In this research, we use 1D CNN instead of 2D CNN. In some cases, CNN can lead to overfitting. In order to prevent overfitting, we also perform dropout. By using dropout, for each training stage, some nodes are either being dropped out or remained in the network using a probability value [10].

Moreover, we also perform PCA before DBN and CNN for some scenarios. We would like to observe the effect of PCA to the performance of the system. PCA can be used to reduce the dimension of the data while also removes noise from the data [1]. PCA can also emphasize strong patterns in the spectra to create better model [11]. We only use PCA to extract features from data and we do not perform any other pre-processing techniques in this research.

3.3. Evaluation
The output of our prediction system is the percentages of macronutrient in a bowl of cooked porridge. However, by using mathematics calculation, we can also infer the needed quantity of the food to reached nutrition intake recommendation for toddlers and children from those percentages.

\[
RMSE = \sqrt{\frac{\sum (y_{pred} - y_{true})^2}{n}}
\]  

(1)

Trained models by CNN and DBN are compared with each other. The evaluation metrics we used are RMSE.
4. Experiments
All instant-porridges are cooked first by using warm water. We use 5 levels of water to capture the effect of water levels in spectral data by reducing about 20 ml water from the ideal water ratio that we get from the package. In this research, we assume that the nutrition facts are valid. After scanning the porridges, we then download the raw spectral data from SCiO Lab web application. In this research, we only try to predict carbohydrate, protein and fat as some of the important nutrition for toddlers beside micronutrient.

4.1. Splitting Dataset
We use 360 data for training and 80 data for testing. In this research, we also use k-fold cross validation with $k=5$ to evaluate the best models. The mean percentages of carbohydrate, protein and fat in training data are 0.1685, 0.0310 and 0.0191 and for testing data 0.1692, 0.0306 and 0.0195 respectively. The carbohydrate percentages range from 0.1 to 0.33, the protein percentages range from 0.016 to 0.075 and fat percentages range from 0.007 to 0.05.

4.2. Models and Parameters
In order to achieve the best model, we observe some scenarios and evaluate the result. We observe DBN and CNN architecture, range values of principal components in PCA and some values of dropout.

4.3. Architecture Observations
In this observation, we evaluate the performance of CNN and DBN. Both CNN and DBN use Adam as optimizer and the number of epochs is 500.

Figure 2 presents the result of this observations with validation data. CNN C achieves best result compared to DBN.

![Figure 2. Macronutrient RMSE in various architectures to validation data](image)

CNN model consists of three blocks where a block contains two convolutional layers and one max-pooling layer. The first block uses 16 3x1 filter, the second uses 32 3x1 filter and the last uses 64 3x1 filters for convolutional layer and 2x1 filter for pooling layer. All convolutional and pooling layers use 2 strides. DBN uses 4 RBM. CNN yields better result compared to DBN due to its convolutional layer that can extract feature from the input data. While processing signal data, it is important to extract feature from the input to make more representative input.

4.4. PCA Observations
We evaluate number of principal components ranging from 40 to 60 to the best model from CNN and DBN.
Figure 3 and Figure 4 reports the result of this observation with validation data. PCA successfully improve the performance for both models. The best principal components value for CNN is 43 and 44 for DBN. CNN still achieves better result compared to DBN.

**Figure 3.** RMSE prediction with and without PCA of CNN to validation data

**Figure 4.** RMSE prediction with and without PCA of DBN to validation data

### 4.5. Dropout Observations

PCA successfully improves the performance of CNN but leads to overfit. In order to prevent overfitting, we try to use dropout with the observed values are 0.1 and 0.2.

Fig. 5 reports the result after using dropout after each convolutional layer and fully-connected layer to validation data. Unfortunately, dropout does not improve the result. Dropout can decrease the training performance but does not improve the validation performance. We suspect that dropout can achieve better performance in larger neural network as in this research, the neural network is quite simple and by adding dropout will make the network smaller and leads to worse performance. We conclude that another way to improve the result is by adding more training data.

**Figure 5.** CNN RMSE Prediction with and without Dropout to validation data

### 5. Discussions and Future Work

We present the use of NIRS and deep learning specifically DBN and CNN to build a macronutrient prediction system of baby food. The experimental results show that CNN achieves best result with prediction errors of carbohydrate, protein and fat achieved by CNN and PCA are 11.70%, 26.14% and 28.72% respectively. This experimental result also means that the correlation between spectral input and macronutrient values is non-linear. However, in this research we can only able to obtain the percentage of macronutrients in the food. In order to achieve more specific and exact result, user needs
to provide the weights of the porridge as the input. We suggest the use of camera combines with SCiO as the camera can measure the volume of the product so the system can give exact and specific result of the macronutrient content in the food.

References

[1] Thong Y J, Nguyen T, Zhang Q, Karunanithi M and Yu L 2017 Predicting food nutrition facts using pocket-size near-infrared sensor 2017 39th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC) pp 742-5

[2] WHO 2017 World Health Organization Accessed April 4th 2018 http://www.who.int/mediacentre/factsheets/malnutrition/en/

[3] BBC 2018 BBC Indonesia Accessed April 4th 2018 http://www.bbc.com/indonesia/indonesia-42872190

[4] Juan X L and Ying Y B 2009 Use of near-infrared spectroscopy and least-squares support vector machine to determine quality change of tomato juice J. of Zhejiang University SCIENCE B 10 (6) pp 465-71

[5] Hailong W, Peng J, Xie C, Bao Y and He Y 2015 Fruit Quality Evaluation Using Spectroscopy Technology: A Review Sensors 15 (5) pp 11889-927

[6] Deng L and Yu D 2014 Deep learning: methods and applications Foundations and Trends® in Signal Processing 7 (3–4) pp 197-387

[7] Huihua Y, Hu B, Pan X, Yan S and Feng Y 2017 Deep belief network-based drug identification using near infrared spectroscopy J. of Innovative Optical Health Sciences 10 (2) 1630011

[8] Asifullah K, Zameer A, Jamal T and Raza A 2018 Deep Belief Networks Based Feature Generation and Regression for Predicting Wind Power arXiv preprint arXiv:1807.11682

[9] Acquarelli J, van Laarhoven T, Gerretzen J, Tran T N, Buydens L M and Marchiori E 2017 Convolutional neural networks for vibrational spectroscopic data analysis Analytica chimica acta 954 pp 22-31

[10] Srivastava N, Hinton G, Krizhevsky A, Sutskever I and Salakhutdinov R 2014 Dropout: a simple way to prevent neural networks from overfitting The J. of machine learning research 15 (1) pp 1929-58

[11] Jonathon S 2014 A Tutorial on Principal Component Analysis arXiv preprint arXiv:1404.1100

[12] Huihua Y, Hu B, Pan X, Yan S and Feng Y 2017 Deep belief network-based drug identification using near infrared spectroscopy J. of Innovative Optical Health Sciences 10 (2) 1630011

[13] Wenwen L, Lin M, Huang Y, Liu H and Zhou X 2017 Near infrared spectroscopy detection of the content of wheat based on improved deep belief network J. of Physics: Conf. Series 887 (1) 012046 (Bristol: IOP Publishing)

[14] Lengkey, Lady C E C H, Budiastra I W, Seminar K B and Purwoko B S 2013 Model pendugaan kandungan air, lemak dan asam lemak bebas pada tiga provenan biji jarak pagar (Jatropha Curcas l) menggunakan spektroskopi inframerah dekat dengan metode partial least square (PLS) Jurnal Penelitian Tanaman Industri 19 (4) pp 203 - 11

[15] Acquarelli J, van Laarhoven T, Gerretzen J, Tran T N, Buydens L M and Marchiori E 2017 Convolutional neural networks for vibrational spectroscopic data analysis Analytica chimica acta 954 pp 22-31