Article

Accurate and Reliable Food Nutrition Estimation Based on Uncertainty-Driven Deep Learning Model

DaeHan Ahn

Department of Electrical, Electronic and Computer Engineering, University of Ulsan,
Ulsan 44610, Republic of Korea; daehani4@gmail.com; Tel.: +82-52-259-2831

Abstract: Mobile Near-Infrared Spectroscopy (NIR) devices are increasingly being used to estimate food nutrients, offering substantial benefits to individuals with diabetes and obesity, who are particularly sensitive to food intake. However, most existing solutions prioritize accuracy, often neglecting to ensure reliability. This oversight can endanger individuals sensitive to specific foods, as it may lead to significant errors in nutrient estimation. To address these issues, we propose an accurate and reliable food nutrient prediction model. Our model introduces a loss function designed to minimize prediction errors by leveraging the relationships among food nutrients. Additionally, we developed a method that enables the model to autonomously estimate its own uncertainty based on the loss, reducing the risk to users. Comparative experiments demonstrate that our model achieves superior performance, with an R^2 value of 0.98 and an RMSE of 0.40, reflecting a 5–15% improvement over other models. The autonomous result rejection mechanism showing a 40.6% improvement further enhances robustness, particularly in handling uncertain predictions. These findings highlight the potential of our approach for precise and trustworthy nutritional assessments in real-world applications.

Keywords: food nutrient estimation; uncertainty reduction; mobile NIRs; mobile healthcare; diabetes and obesity

1. Introduction

Over the past decade, there has been a significant increase in food-related health issues, such as obesity and diabetes [1–3], affecting billions of the global population [4–6]. To help people manage these conditions, self-management is becoming increasingly important for maintaining health and wellness. The Centers for Disease Control and Prevention (CDC) recommend that individuals record their intakes of daily calorie and nutrients, such as carbohydrates, proteins, and fats (CPF), which are essential for managing weight and blood sugar levels [7].

To facilitate the recording of food intake for individuals, various mobile applications have been proposed to estimate the calorie (Cal) and CPF contents of meals [8–15]. Most of these apps use a smartphone camera to visually assess the food and estimate the nutrients based on image recognition technology [16–20]. While the camera-based approaches offer convenience and accessibility, they often struggle to provide accurate estimations for cooked food or meals with complex ingredients. To address the limitations, alternative approaches using near-infrared spectroscopy (NIR) have gained attention. The NIR identifies macronutrients based on the absorption characteristics of the chemical bonds (C-H, O-H and N-H) in the near-infrared region. Each macronutrient (CPF) has unique absorption peaks at specific wavelengths in the NIR spectrum. This makes NIR a promising solution for analyzing a wide variety of foods [21–25]. In particular, deep learning models based on a supervised approach excel at generating proposed outputs from specific inputs, making them highly effective for analyzing Cal and CPF values from the NIR signal of food.

Most studies using NIR cameras [26–32] have focused on developing regression models to predict nutrient values such as Cal and CPF. However, these approaches often overlook the uncertainty associated with the reliability of the prediction result. In this case, for individuals who are highly sensitive to specific foods, even minor inaccuracies in calorie estimation can lead to severe consequences. Therefore, it is essential to reduce the uncertainty in the prediction model to ensure their safety. Additionally, a notification system is necessary to inform users when a prediction result may be unreliable due to high uncertainty, enabling them to decide whether to trust the estimate.

In this paper, we propose a novel approach to food nutrient prediction that aims to reduce model uncertainty and improve prediction accuracy. The proposed model achieves these improvements by leveraging the natural relationship between Cal and CPF values, which are related by the fact that their values should ideally be equal. As prediction errors cause deviations between these values, this difference forms the basis for a reliability metric that allows for the model to autonomously discard unreliable predictions. Based on that, we introduce a metric for measuring the uncertainty of prediction based on this deviation and formulate an uncertainty-driven objective function designed to minimize the uncertainty and improve the accuracy. Furthermore, the proposed objective function does not require additional parameters other than those needed for food nutrient prediction, thereby eliminating the need for extra architectures or multimodal layers. Consequently, our approach enables the development of a simpler and more efficient model.

In summary, our main contributions are as follows:

- Uncertainty-Driven Prediction and Rejection. We introduce an uncertainty assessment metric that allows for the model to autonomously reject predictions deemed unreliable due to high uncertainty.
- Object Function for Uncertainty Reduction. We propose a specialized loss function that reduces uncertainty in predictions by focusing more on reliable data and less on uncertain data during training, which enhances model performance and robustness.

2. Methods

2.1. Model Architecture

The process to estimate Cal and CPF values from NIR signals can be defined as follows. Let $X \in \mathbb{R}^{1 \times D}$ be the input, where X represents the NIR signals with a sensing range of 256 (i.e., $D = 256$) for a given food item. The output Y comprises nutrient values $\{Calorie, Carbohydrate, Protein, Fat\}$ corresponding to the input X. The deep learning model is denoted by $f_\theta()$, where θ represents the model's training parameters. The primary objective of the model is to map the input to the output, expressed as $Y = f_\theta(X)$. To achieve this objective, we designed the model architecture illustrated in Figure 1. The model consists of consecutive hidden layers that form the overall architecture. Each hidden layer is denoted by $\sigma(z)$, where σ is the activation function, and z is the output from the previous layer. Each layer employs the Rectified Linear Unit (ReLU) function as the activation function, due to its widespread use in non-linear models. This model simply predicts the output corresponding to the given input, and it does not require other structural requirements, allowing for the creation of a simple network. During the prediction phase, the model utilizes our proposed uncertainty criterion to autonomously decide whether to accept or reject the result. Further details on this decision-making process are provided in Sections 2.2 and 2.3. During training, the model adapts to a specifically designed loss function aimed at estimating uncertainty. The specifics of this loss function are detailed in Section 2.4. This tailored approach ensures that the model not only predicts nutrient values with high accuracy but also assesses the reliability of its predictions, enhancing overall model robustness and performance.
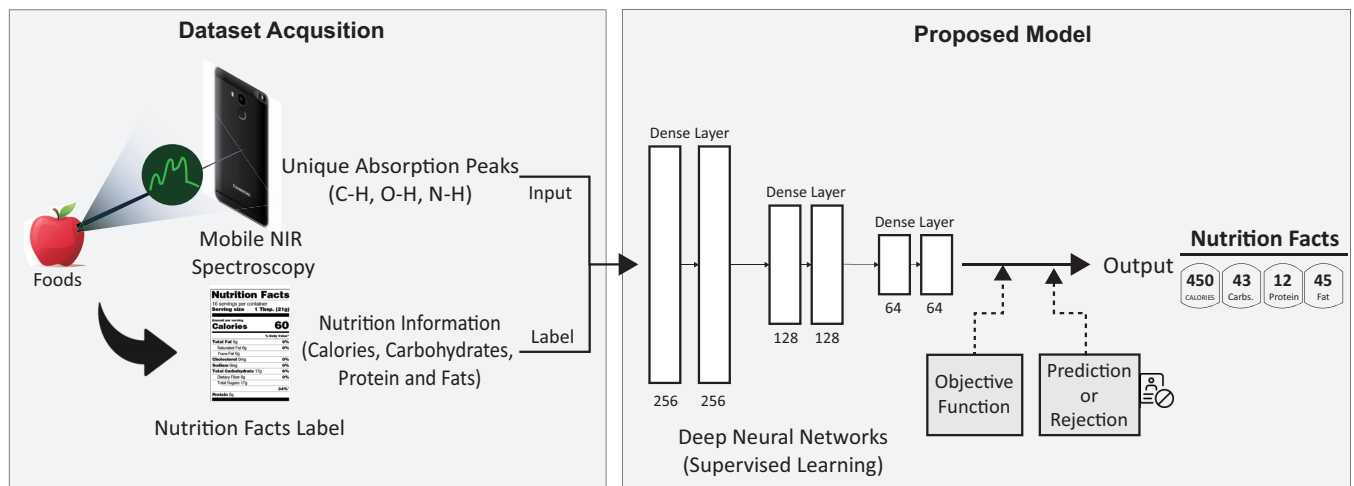
Figure 1. Overview of the proposed system. The system combines NIR spectroscopy with an uncertainty-aware deep learning model to estimate calories and macronutrient contents (carbohydrates, proteins, and fats) in food. It analyzes NIR absorption patterns for accurate nutrient prediction and flags uncertain results to improve reliability.

2.2. Assessment of Uncertainty

The proposed model is designed to estimate both *Cal* and *CPF* values. Under ideal conditions, the optimal values estimated by a model, denoted as Cal^* and CPF^*, are characterized by the following relationship:

$$Cal^* = CPF^*, \tag{1}$$

$$CPF^* = (4 \times Carbohydrate^*) + (4 \times Protein^*) + (9 \times Fat^*) \tag{2}$$

The equations indicate that the difference between Cal^* and CPF^* should be zero. However, in the real case, prediction errors are unavoidable. Moreover, due to the independent prediction paths for Cal and CPF, discrepancies can arise between these two values, despite their theoretical equivalence. Based on this motivation, we observed how the difference between the relationships varies according to the prediction performance based on the deep neural network model illustrated in Figure 1. Figure 2 displays the boxplot of errors, measured using the Symmetric Mean Absolute Percentage Error (SMAPE), between the predicted calorie values (*Cal*) and the calorie values (*Cal′*) estimated from the predicted CPF values for the model trained to predict food components. The figure highlights both the best 30% of cases, where the model achieved the highest accuracy, and the worst 30% of cases, where the largest prediction errors occurred. The results reveal that the best and worst cases have median SMAPE values of 0.07 and 0.15, respectively. Furthermore, the upper whiskers, representing the 95th percentile, are 0.19 and 0.55, respectively. The outliers beyond the upper whisker are plotted using the '+' symbol. These findings indicate a correlation where the SMAPE between *Cal* and *Cal′* increases proportionally with the augmentation of estimation error, highlighting the direct impact of the error magnitudes on the model's performance metrics. In other words, this experiment demonstrates that the difference between the *Cal* and *Cal′* values can be used to measure the model's uncertainty. Analyzing these discrepancies provides insights into the reliability of the model's predictions, thereby offering a metric for assessing its performance and robustness in practical applications.

Motivated by these findings, we propose a metric to estimate the uncertainty, denoted as U_i, as follows:

$$U_i = \frac{|\, Cal_i - Cal_i' \,|}{(|\, Cal_i \,| + |\, Cal_i' \,|)/2} \tag{3}$$

where i denotes the index of each sample. Cal is the calorie value estimated by a model, and Cal' is the calorie value estimated from the predicted CPF values. The uncertainty values U_i, ranging from 0.0 to 1.0, represent the global lower and upper bounds of the uncertainty function. A value of 0.0 indicates perfect prediction, while 1.0 denotes maximum uncertainty. These bounds are inherent to the metric and remain independent of the dataset, though the observed distribution of values may vary based on the specific data. Consequently, this uncertainty metric enables the model to autonomously estimate its own error. This self-assessment capability enhances the model's reliability and robustness in practical applications by providing a quantifiable measure of prediction confidence.
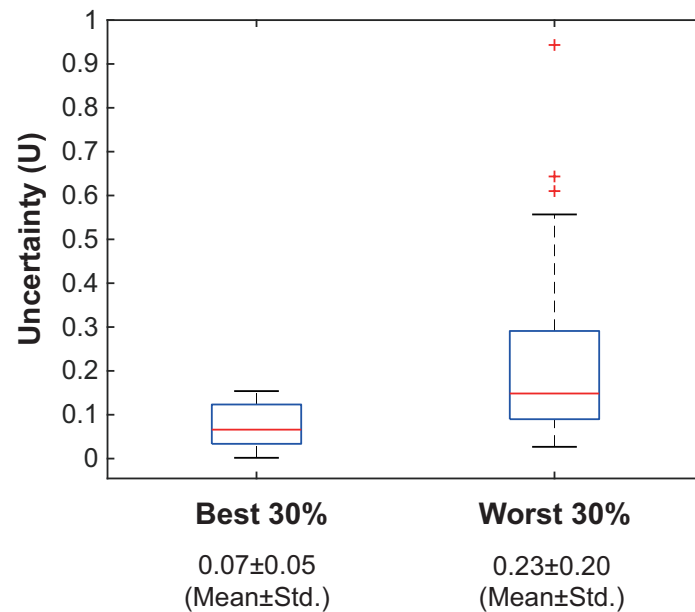

Figure 2. Boxplot comparison of uncertainty (U) between the best 30% and worst 30% of cases in predictions. The worst group shows greater uncertainty, with a wider interquartile range and several outliers, marked by red crosses, highlighting the group's variability.

2.3. Rejection of Unreliable Result

The uncertainty assessment metric introduced in Section 2.2 ranges from 0.0 to 1.0 and indicates a high value when the model predicts that the estimation result might be unreliable. Consequently, this uncertainty can effectively serve as a confidence score, enabling the model to autonomously determine the reliability of its outcomes. With this score, the model possesses the capability to autonomously reject results it deems poor, enhancing its robustness and reliability.

Additionally, the confidence score can serve as valuable information for users. If the model displays this score, it provides users with a quantifiable measure of reliability, allowing them to make informed decisions about whether to rely on the predicted data. For instance, in the context of nutritional assessments, users can decide whether a particular food intake is advisable based on the estimated nutritional content and its associated confidence score. This added transparency can significantly improve user trust and satisfaction, as they are better informed about the certainty of the model's predictions.

Overall, the integration of an uncertainty-based confidence score not only improves the model's decision-making process but also enhances user interaction by providing a clear and interpretable measure of prediction reliability. This dual benefit underscores the importance of effectively managing uncertainty in predictive modeling, particularly in applications requiring high precision and user trust.

2.4. Objective Function for Uncertainty Reduction

The Figure 3 illustrates the distribution of calorie values within our data sample and highlights the worst 30% of predictions on the distribution. The figure show which types of data the model finds difficult to identify. The results indicate that there are almost no worst predictions for data with a large number of samples. In contrast, for data with fewer samples, specifically those at the ends of the data distribution, the prediction performance is lower. This pattern suggests a link to epistemic uncertainty [33], which may arise from an insufficient amount of data available for effective learning. According to a study [33], such uncertain data can prevent the model's ability to be properly regularized to the observed data during training, thereby affecting its performance and reliability. To address this issue and maximize learning efficiency, we propose a loss function that focuses more on reliable data using an uncertainty criterion while having less interest in data with high uncertainty. The proposed the loss function is defined as follows in Equation (4).

$$\mathcal{L}(\theta, y) = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{|y_i - f_\theta(x_i)|^2}{max[1, \alpha \cdot exp(U_i)]} + \beta \cdot [log(U_i) + 1] \right) \tag{4}$$
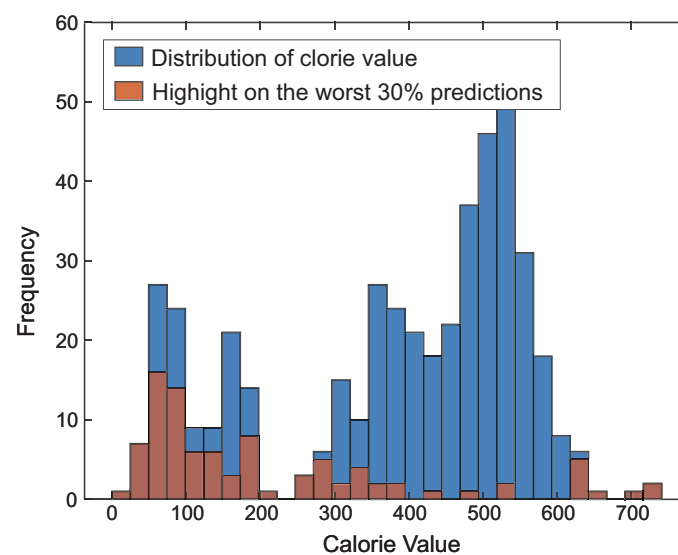

Figure 3. Distribution of calorie value in the dataset, with the worst 30% of predictions highlighted. The model performs worse on data with fewer samples, particularly at the distribution extremes.

The proposed loss function is composed of two components. The first component is a residual term that calculates the difference between the actual value y_i and the predicted values $f_\theta(x_i)$. The uncertainty, expressed as $exp(U_i)$, is applied as a weight to adjust the model's strictness, allowing for the model to focus on data with low uncertainty during the training process. For instance, when uncertainty is high, the residual error is intentionally reduced below the actual value to promote weaker learning. Conversely, when uncertainty is low, the residual error reflects the actual value more closely, encouraging more focused learning. Additionally, there is a constant α in the function to adjust its strictness, which can result in values less than 1.0. This issue is addressed by using the max function. The second component is regularization, which aims to decrease uncertainty during training and to improve the estimation performance. In this context, β is a constant with a small value (≤ 0.1). As a result of that, the loss ensures that results with high uncertainty are considered less important during the learning process, while data with high confidence are learned more intensively by the model.

3. Related Works and Baselines

The studies related to the mobile food nutrient analysis can be categorized into a food-intake recommendation system, camera-based vision analysis system, and NIRs-based food analysis system.

A food recommendation system [10,15] helps users plan meals by suggesting meal timings and diets based on their health status. By entering details like height and weight, it generates a customized diet plan and recommends food within a specific calorie range. These systems are particularly useful for managing conditions like obesity and diabetes through personalized nutrition advice. Other studies [11,14] have developed systems that use IoT technology to accurately monitor food intake, allowing for more precise diet management by tracking what is actually consumed. Recent studies [9,16] have integrated machine learning algorithms to enhance the personalization and accuracy of these recommendations by analyzing health data and eating patterns. This ensures that dietary advice adjusts to the user's changing health needs and preferences.

Camera-based methods [15,34,35] are the most widely researched approach to food recognition and calorie estimation. This approach involves taking photos of food with a mobile camera, identifying the food item, and displaying its calorie content based on data retrieved from a food database. This approach has gained particular popularity due to their simplicity and convenience. From a user-experience perspective, it is easy to capture and log the nutritional content of a meal just by snapping a photo with one's own smartphone. Additionally, the RGB-based method is relatively straightforward to develop, using image recognition technology to identify foods, connect to a dataset, and pull up the relevant nutritional information. However, despite being easy to use, this method struggles with accurately estimating the nutritional values of cooked food and meals with different recipes. This is mainly due to the limitations of the database, which only provides data for the stored food names and cannot account for foods not already in the system.

To address these limitations of RGB cameras, applications using mobile NIR technology are being developed. NIR cameras can capture the chemical characteristics derived from nutritional bonds, making them more accurate for estimating nutrients in a wide variety of foods [21–24]. Most studies using NIR camera [21,22,26–32] have focused on using prediction models to estimate values such as Cal and CPF, but they often overlook the uncertainty in these predictions. This uncertainty is crucial for individuals who are highly sensitive to specific foods as even small inaccuracies can have serious consequences. To tackle this, a previous study [31] proposed a neural network model with a multi-modal structure to reduce uncertainty. The proposed model consists of the two components that are the reconstruction for analyzing the food's NIR image and the prediction for estimating the nutrient values. The idea is that if a major error occurs in analyzing the food's NIR image, the predicted nutrient values cannot be trusted either, as both predictions are connected in the root layer of the neural network. Therefore, the study used errors in the NIR image analysis to reject unreliable nutrient predictions. However, this approach struggles to balance learning between analyzing the food's NIR image and predicting the nutrients. The reconstruction component requires a significantly larger number of terminal nodes compared to those needed for nutrient estimation. Consequently, during the training phase, the accumulated loss from the reconstruction nodes surpasses the loss from the nutrient estimation nodes. This imbalance can lead the model to be more sensitive to the reconstruction process than the estimation of the nutrient values [36,37].

We identified reference models that work in a similar environment to ours. Our environment involves using a portable NIR device to capture 1D NIR signals, which are then processed to predict the nutritional content of food. One of the most comparable models is NIRSCam [30], which uses several traditional machine learning methods such as K-Nearest Neighbor (KNN), logistic regression (LR), decision tree (DT), random forest (RT), and XGboost (XGB) [38–42] to analyze food composition. Additionally, we incorporate two deep learning-based models: CNN [28] and Multi-Modal [31]. The CNN model utilizes 1D convolution layers to analyze food components, while the Multi-Modal (MM) model com-

prises two distinct components that are nutrient-prediction nodes and reconstruction nodes. The reconstruction nodes are responsible for reconstructing the provided input, and the model calculates the difference between the input and the reconstructed value to determine whether to reject the result based on this difference. We conducted evaluations using the model denoted as MM, which employs the estimated results from the given model without any additional post-processing. Additionally, we named and evaluated the optimized model as MM*, which removes uncertain results from the estimated outputs. Similarly, for our proposed model, we performed evaluations by distinguishing between the Proposed model and the optimized Proposed model as Proposed and Proposed* respectively.

While there are other models available, they either handle different types of data or share similar characteristics with the representative models mentioned above. Therefore, we chose these models as representative baselines for our evaluation. Furthermore, the Multi-Modal model was chosen as a comparative target for reliability assessment, as it includes a mechanism to discard uncertain predictions, which aligns with our proposed idea. This comparison ensures a robust comparison of reliability and performance across various modeling approaches.

Recent studies [43–45] have explored sensor fusion approaches that incorporate RGB, NIR, and other modalities for food classification and caloric estimation, achieving significant success in nutrient analysis. While these methods rely on features captured through multispectral imaging, combining both visual and near-infrared data, our approach focuses exclusively on NIR spectroscopy. By integrating an uncertainty-driven deep learning model, we aim to improve both the accuracy and reliability of nutrient estimation, distinguishing our solution from others that solely pursue accuracy. Additionally, most models utilize a deep learning architecture based on CNNs, which is closely related to the architecture presented in the study [28]. Therefore, we evaluated our NIR-based analysis using a CNN framework, with [28] considered as a reference. Sensor fusion approaches, however, remain a promising solution for enhancing accuracy and may continue to gain attention for their ability to leverage diverse data sources effectively. Our proposed method has the potential to further contribute to advancements in this field by integrating sensor fusion solutions in future iterations.

4. Evaluation Methods

4.1. Datasets

We employed the food nutrition datasets introduced by the referenced study [31]. These datasets comprise detailed nutritional information on over 400 food items, including a variety of food groups such as dairy products, meats, fruits, and instant foods. Each data entry includes comprehensive nutrient facts such as Cal and CPF values, alongside the corresponding NIR signal. The NIR signals were captured using a mobile NIR device [27] that operates within a sensing range of 780–1070 nm and utilizes 256 frequency bands to ensure precise measurements. We utilized these data pairs, where the NIR signal serves as the input and the nutrient facts as the output for the model. This approach allows for a comprehensive assessment of the model's ability to accurately predict nutritional content based on NIR signals. The datasets employed in this study are available from the corresponding authors [31] upon reasonable request.

4.2. Experimental Setup

The experimental methodology employed in this study involves a 10-fold cross-validation process. Specifically, the dataset is divided into 10 equal-sized sub-datasets. In each of the 10 folds, 9 subsets are used for training the model, while the remaining subset is reserved for testing. This procedure is repeated 10 times, with each subset serving as the test set in one fold and as part of the training set in the other nine folds. At the end of these cycles, the results from all 10 folds are aggregated to provide a thorough evaluation of the model's performance. This approach not only ensures the reliability of the model by

testing it across different subsets but also helps in assessing the consistency of its predictive accuracy across various segments of data.

The food nutrients evaluated in the experiments include Cal and CPF, totaling four components. Each of these components has a different scale. To ensure consistent evaluation across all components, we normalized the data using Z-score normalization. This method adjusts the data such that all features have a mean of 0 and a standard deviation of 1, conforming to the properties of a standard normal distribution.

4.3. Threshold Optimization

The proposed model incorporates a threshold to recognize uncertainty. If the model's assessed uncertainty is below the threshold, the result is accepted; if it exceeds the threshold, the result is rejected. This characteristic introduces a trade-off between performance and the rejection rate based on the threshold setting. For example, Figure 4 depicts the preliminary results according to the threshold value. As shown in Figure 4a, a high threshold accepts most results and allows most poor predictions. Conversely, as shown in Figure 4b, a low threshold rejects most results including many poor predictions. However, too frequent rejections of results can hinder the practical usability. Therefore, determining an appropriate threshold is crucial for accurately evaluating the model's performance. The optimal values is the one that maximizes performance while minimizing rejections.
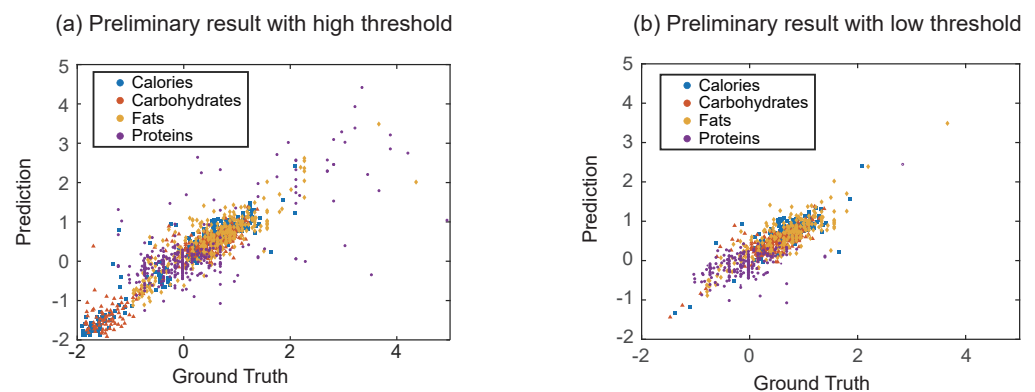

Figure 4. Preliminary results with different uncertainty thresholds. (**a**) High threshold: accepts most predictions, including poor ones. (**b**) Low threshold: rejects poor predictions. This highlights the trade-off between performance and rejection rate.

To address this issue, we approached the problem statistically. Figure 5b depicts the p-value representing the difference between *Cal* and *CPF* predictions relative to the rejection threshold. At high thresholds, the distributions of the two predictions are nearly identical, but as the threshold decreases, a statistical difference emerges. When the p-value falls below a certain level (<0.05), it supports the alternative hypothesis that a significant difference exists. This implies that using Equation (3) to calculate uncertainty is ineffective at these levels. To balance acceptance and reliability, we set the significance level at p-value ($=0.05$) to determine the threshold, thereby supporting the null hypothesis that there is no statistically significant difference between the distributions. This approach ensures that uncertainty is assessed using Equation (3) at a statistically meaningful level.
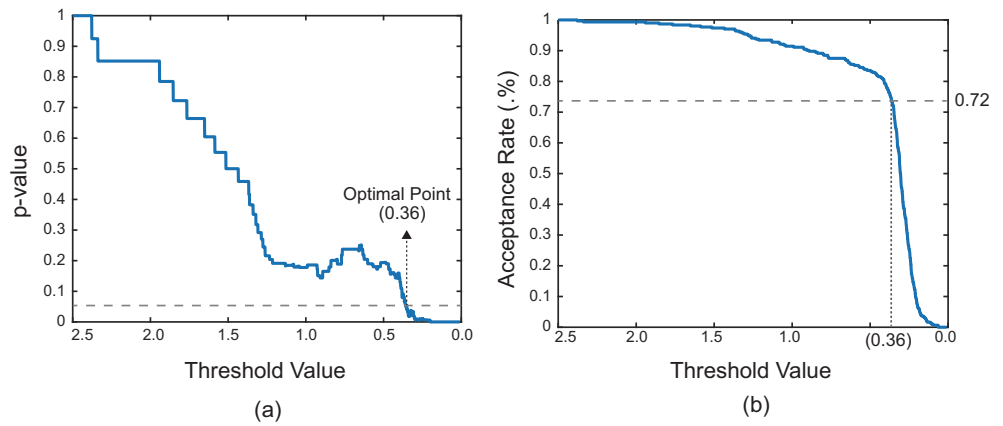
Figure 5. Statistical analysis for determining the optimal threshold for balancing prediction acceptance and reliability. (**a**) Relationship between *p*-value and threshold, with 0.36 identified as the optimal threshold at a *p*-value of 0.05. (**b**) The acceptance rate as a function of threshold, where 0.36 corresponds to a 72% acceptance rate.

When setting the threshold to a *p*-value of 5%, a value of 0.36 can be obtained, and when applied to Figure 5a, it shows that the model accepts the prediction results at a 72% level.

4.4. Performance Evaluation

The proposed model is a regression model designed to predict numerical values and is evaluated using two metrics. The first metric is the adjusted coefficient of determination, which is one of the most well-known metrics for evaluating regression models. This metric indicates how well the model explains the variance of the dependent variable and is expressed as follows:

$$R^2_{adj} = 1 - \frac{(1 - R^2)(n - 1)}{n - k - 1} \tag{5}$$

$$R^2 = \frac{\sum(y_i - \bar{y})^2}{\sum(y_i - \hat{y}_i)^2} \tag{6}$$

where y_i and $\bar{y}$ are the target value and mean of the dataset, and $\hat{y}$ is a prediction. n and k are the number of observation and predictors.

The second metric used is the root mean square error (RMSE), which is expressed as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum(y_i - \hat{y}_i)^2} \tag{7}$$

This metric represents the average squared difference between the actual values and the values predicted by the model, indicating how close the model's predictions are to the actual data.

By using these two evaluation metrics, we aim to directly assess the model's explanatory power for the dependent variable and the prediction error.

4.5. Assessing the Effectiveness of Reduction

While performance metrics are important for evaluating the proposed model, it is even more crucial to verify the effectiveness of the proposed rejection metric. The proposed model is designed to reduce uncertainty, and it identifies and removes the worst predictions based on the uncertainty metric before producing the final output. Therefore, it is necessary to evaluate by how much the worst predictions have been reduced. To measure this, we set the percentage error and Z-score between the actual values and the prediction as the

evaluation metrics. The percentage error is a measure of the deviation of a measurement from the actual value, expressed as a percentage. The formula is as follows:

$$\text{Percentage Error} = \frac{y_i - \hat{y}_i}{y_i} \times 100 \tag{8}$$

This metric has the advantage of expressing both overestimation and underestimation in a regression model. Thus, it is useful for providing insights into the measurement's inaccuracy and error distribution. We also adopt the Z-score to evaluate the models in the effectiveness of the uncertainty reduction. The Z-score provides a statistical measurement that describes the relative error from the mean of a group. It is calculated using the following formula:

$$\text{Z-Score} = \frac{1}{n} \sum \frac{\hat{y}_i - \bar{y}}{\sigma} \tag{9}$$

where σ is standard deviation. The Z-score can be used to evaluate by how far the model's predictions deviate from the actual values, thus providing a measure of uncertainty. This method allows for an assessment of the reduction in error through the implementation of uncertainty quantification techniques.

5. Experimental Results

Figure 6 depicts the demonstration results of the baselines and proposed model. In the figure, the closer the points are to the reference line, the better the predictions. All models exhibit an upward trend corresponding to the reference line, demonstrating their effectiveness in food component analysis, though some notable differences exist. When comparing the results of the deep learning-based model in the second column with the traditional machine learning techniques in the first column, the traditional methods are further from the reference line with many predictions significantly deviating. In contrast, the deep learning-based models are relatively closer to the reference line, with a considerable reduction in the number of predictions significantly deviating from the actual values.
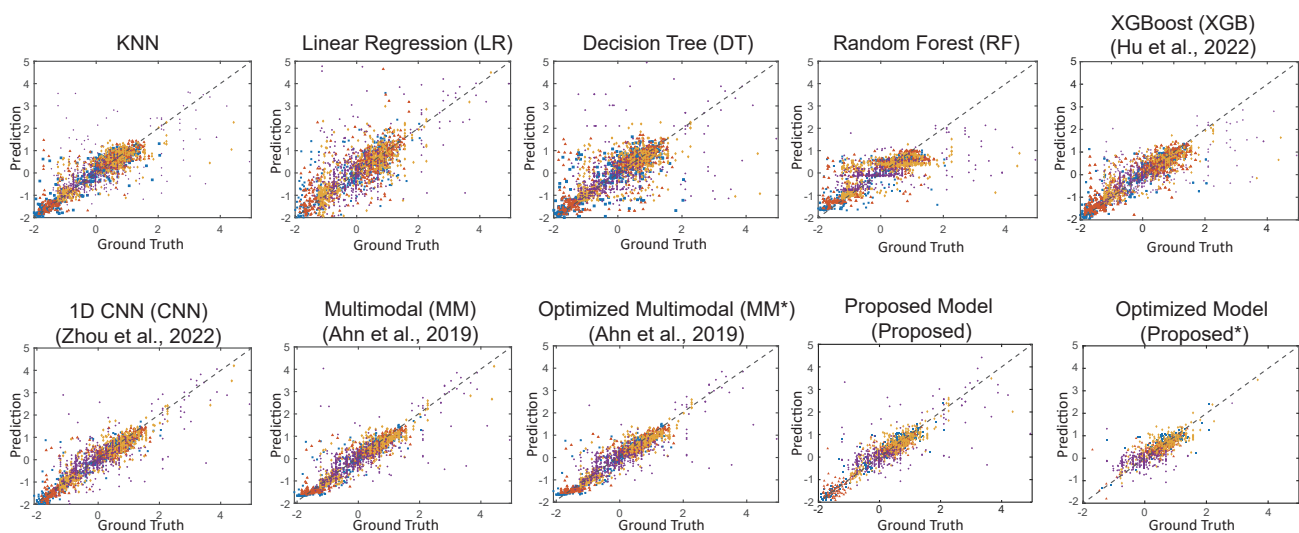

Figure 6. Comparison of the prediction results across baselines and the proposed model. The traditional models (KNN, LR, DT, RF, and XGB) show a greater deviation from the reference line, while deep learning models (on the bottom rows) align more closely, indicating superior performance [28,30,31].

Table 1 provides a quantitative evaluation for the aforementioned experiment. The traditional machine learning models introduced in [28], such as KNN, LR, DT, RF, and XGB, demonstrated varying performance. The XGB model achieved an R^2 of 0.94 and a higher

RMSE of 0.56. On the contrary, the LR model exhibited the lowest R^2 value of 0.87 and the highest RMSE of 0.89, indicating the least accuracy among the models evaluated. The deep learning-based model averaged 0.958 for R^2, which is higher than the 0.912 for traditional models. Among the models evaluated, the proposed* model achieved the highest R^2 value of 0.98. But the CNN, MM, and MM* also demonstrated strong performance with R^2 values ranging from 0.94 to 0.96. In addition, the proposed* model showed the lowest RMSE of 0.40, achieving an improvement in RMSE of 5 to 15% over the other deep learning based-models. The results show that incorporating our proposed loss function and rejection metric significantly enhances model performance.

Table 1. Comparison results in coefficient of determination (R^2) and RMSE. The proposed model achieves the best performance with an R^2 of 0.98 and RMSE of 0.40.

	KNN	LR	DT	RF	XGB [30]	CNN [28]	MM [31]	MM* [31]	Proposed	Proposed*
R^2_{adj}	0.93	0.87	0.90	0.92	0.94	0.95	0.94	0.96	0.96	0.98
RMSE	0.58	0.89	0.78	0.64	0.56	0.46	0.45	0.42	0.44	0.40

Figure 7 illustrates the model's percentage error shown as a boxplot. This visualizes the overall distribution of errors, with extreme errors denoted by the '+' symbol. In evaluating these distributions, a model with short tails is considered effective. All models exhibit an average percentage error close to zero, which reflects good average performance in food nutrient analysis. This observation is corroborated by Table 2, which shows that the average error is close to zero, demonstrating the accuracy of the models. However, a detailed examination reveals significant differences in the error distributions among the models. The tails of the traditional models are generally longer than those of the deep learning models, indicating that traditional models tend to produce more extreme errors. This variability implies that traditional models are less consistent in their predictions. In contrast, the deep learning models show a short tail, meaning that errors occur within a smaller range. This consistency is indicative of the robustness of deep learning models. Among these, the proposed model displays the smallest range of errors, highlighting its precision. Furthermore, the optimized model (proposed*), which includes an autonomous result rejection mechanism, narrows the maximum error range to align closely with the confidence interval of the deep learning model, demonstrating enhanced reliability.
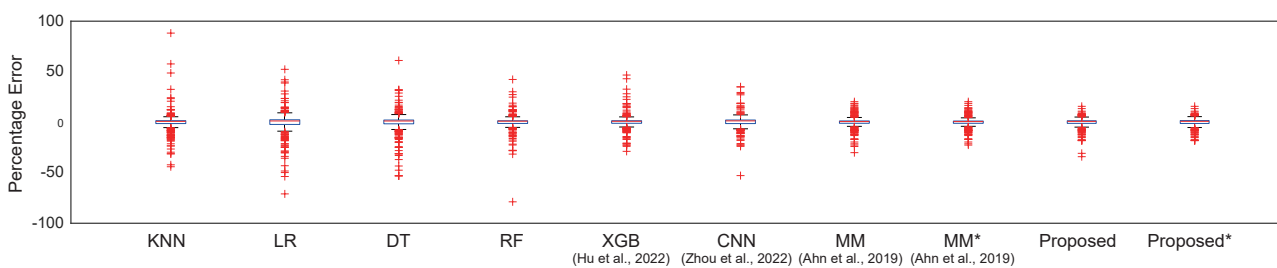

Figure 7. Boxplot comparison for the percentage error. The traditional models (KNN, LR, DT, RF, and XGB) have higher variability, while deep learning models (CNN, MM, proposed) are more consistent. The optimized proposed model has the smallest error range, indicating its precision and reliability [28,30,31].

Table 2. Comparison of the baselines and the proposed model in terms of mean and Z-score. The proposed model achieves the lowest mean error (0.24) and Z-score (0.64), indicating superior performance.

	KNN	LR	DT	RF	XGB [29]	CNN [28]	MM [31]	MM* [31]	Proposed	Proposed*
Mean	0.34	0.57	0.46	0.41	0.35	0.28	0.26	0.24	0.26	0.24
Z-score	0.91	1.24	0.97	0.92	0.87	0.83	0.85	0.80	0.82	0.64
Z-score (worst 30%)	1.42	2.06	1.52	1.44	1.30	1.23	1.25	1.20	1.21	0.86

The quantitative experimental results are comprehensively presented in Table 2. We evaluated the various models against the mean error, the Z-score, and the Z-score for the worst 30% of predictions. As observed in Table 1, the deep learning-based models, particularly the proposed models, demonstrated superior mean performance with values of 0.24–0.26. This represents a significant improvement over traditional machine learning models, which exhibit higher mean errors of 0.34–0.57. These results highlight the effectiveness of deep learning approaches in reducing average prediction errors in food component analysis.

The Z-score measures the standard deviation of a prediction, providing insight into the consistency and reliability of the model's predictions. Lower Z-scores indicate less variability and more reliable predictions. Therefore, the effectiveness of the rejection can be measured by this metric. As a result, traditional models such as KNN, LR, DT, RF, and XGB achieved Z-scores of 0.91, 1.24, 0.97, 0.92, and 0.87, indicating greater variability and less reliability in their predictions. The comparative analysis of deep learning models versus traditional machine learning models reveals that the deep learning models outperform traditional approaches. For instance, CNN, MM, and MM* achieved an average mean error of 0.26 and a Z-score of 0.82; they are superior to those of traditional models. This result highlights the robustness of the deep learning models in maintaining a consistent performance. Moreover, the evaluation of the proposed model shows that it had a mean error of 0.26 and a Z-score of 0.82; the optimized model (proposed*) demonstrated further improvements with a mean error of 0.24 and a Z-score of 0.64. This indicates that the inclusion of an autonomous result rejection mechanism significantly enhances model performance.

The Z-score for the worst 30% of predictions is a critical metric that evaluates the model's performance in handling the most challenging cases. It reflects the model's ability to minimize the impact of uncertain predictions. The proposed model reduced the Z-score (worst 30%) from 1.21 to 0.86 after implementing the result rejection method, showing a 40.6% improvement. These results are improvements to the level of the mean Z-score of the existing deep learning models. Additionally, unreasonable predictions are completely eliminated. It achieved a remarkable performance compared to the evaluation result observed in MM* (=1.20). This dramatic reduction in error for the worst predictions confirms the efficacy of our approach in enhancing model reliability and accuracy.

6. Discussion

This study adopted and validated a deep learning model with a simple structure, as we observed that employing modern models with a large number of parameters could easily lead to overfitting. Consequently, it was determined that a shallow network with fewer parameters would be most effective for the given datasets. In the future, as the size of the data increases, it may be feasible to use larger models. In such cases, the proposed loss function could be employed to enhance performance.

The proposed system is based on NIRs. This approach has the advantage of collecting data closely related to food composition. However, it faces challenges, such as the difficulty of developing mobile-sized hardware and the limitation of scanning only localized areas of the sample. To make this system practically useful, a more integrated approach is required. Such a system would leverage NIRs in combination with camera-based visual recognition and user input, among other user-friendly features.

As highlighted in Section 2.2, it is observed that uncertainty does not always align with the magnitude of error. Consequently, the model might reject a result that is relatively accurate due to a high perceived uncertainty. The analysis of Figure 4 indicates that such instances often arise when data samples fall outside the customary distribution range, potentially linked to data sparsity issues. Our future research efforts will focus on addressing concerns related to data sparsity.

7. Conclusions

In this study, we introduce a food nutrient prediction model based on NIR technology. The proposed model leverages a unique loss function that enhances prediction accuracy by minimizing the difference between predicted *Cal* and *CPF* values. Additionally, we incorporate a reliability evaluation metric that allows the model to autonomously reject unreliable results, significantly improving its robustness.

The comparative analysis with traditional machine learning models and other deep learning approaches demonstrated that the proposed model achieved superior performance. With an R^2 value of 0.98 and an RMSE of 0.40, our model showed a 2–11% improvement over other models in R^2 and achieved a 1.05–2.22 times improvement in RMSE. The evaluation also highlighted the importance of the autonomous result rejection mechanism, which effectively reduced the Z-score for the worst 30% of predictions from 1.21 to 0.86, confirming its efficacy in enhancing model reliability and accuracy.

Our proposed model represents a significant advancement in the field of automated dietary assessment, providing accurate and reliable nutritional information that can aid in the management of food-related diseases. This study underscores the potential of combining NIR technology with deep learning to improve health outcomes through precise and trustworthy nutritional assessments.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data underlying this article will be shared upon a reasonable request to the corresponding author.

Conflicts of Interest: The author declares no competing interests.

References

1. Obesity and Overweight. Available online: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight (accessed on 13 March 2024).
2. Chong, B.; Jayabaskaran, J.; Kong, G.; Chan, Y.H.; Chin, Y.H.; Goh, R.; Kannan, S.; Ng, C.H.; Loong, S.; Kueh, M.T.W.; et al. Trends and predictions of malnutrition and obesity in 204 countries and territories: An analysis of the global burden of disease study 2019. *eClinicalMedicine* **2023**, *57*, 101850. [CrossRef]
3. Harding, J.L.; Pavkov, M.E.; Magliano, D.J.; Shaw, J.E.; Gregg, E.W. Global trends in diabetes complications: A review of current evidence. *Diabetologia* **2019**, *62*, 3–16. [CrossRef]
4. Dong, C.; We, G.; Li, H.; Qiao, Y.; Gao, S. Type 1 and type 2 diabetes mortality burden: Predictions for 2030 based on Bayesian age-period-cohort analysis of China and global mortality burden from 1990 to 2019. *J. Diabetes Investig.* **2024**, *15*, 623–633. [CrossRef] [PubMed]
5. Saeedi, P.; Petersohn, I.; Salpea, P.; Malanda, B.; Karuranga, S.; Unwin, N.; Colagiuri, S.; Guariguata, L.; Motala, A.A.; Ogurtsova, K.; et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas. *Diabetes Res. Clin. Pract.* **2019**, *157*, 107843. [CrossRef] [PubMed]
6. Ward, Z.J.; Bleich, S.N.; Cradock, A.L.; Barrett, J.L.; Giles, C.M.; Flax, C.; Long, M.W.; Gortmaker, S.L.; Projected, U.S. State-Level Prevalence of Adult Obesity and Severe Obesity. *N. Engl. J. Med.* **2019**, *381*, 2440–2450. [CrossRef] [PubMed]
7. Centers for Disease Control and Prevention (CDC), Living with Diabetes. Available online: https://www.cdc.gov/diabetes/living-with/index.html (accessed on 13 March 2024).
8. Azumio Inc. Calorie Mama AI: Diet Counter. Available online: https://www.caloriemama.ai/ (accessed on 13 March 2024)
9. Rostami, M.; Farrahi, V.; Ahmadian, S.; Jalali, S.M.J.; Oussalah, M. A novel healthy and time-aware food recommender system using attributed community detection. *Expert Syst. Appl.* **2023**, *221*, 119719. [CrossRef]

10. Wang, X.; Dou, Z.; Feng, S.; Zhang, Y.; Ma, L.; Zou, C.; Bai, Z.; Lakshmanan, P.; Shi, X.; Liu, D.; et al. Global food nutrients analysis reveals alarming gaps and daunting challenges. *Nat. Food* **2023**, *4*, 1007–1017. [CrossRef] [PubMed]

11. Shi, Z.; Li, X.; Shuai, Y.; Lu, Y.; Liu, Q. The development of wearable technologies and their potential for measuring nutrient intake: Towards precision nutrition. *Nutr. Bull.* **2022**, *47*, 388–406. [CrossRef]

12. Stuart, M.B.; McGonigle, A.J.S.; Davies, M.; Hobbs, M.J.; Boone, N.A.; Stanger, L.R.; Zhu, C.; Pering, T.D.; Willmott, J.R. Low-Cost Hyperspectral Imaging with a Smartphone. *J. Imaging* **2021**, *7*, 136. [CrossRef] [PubMed]

13. Chotwanvirat, P.; Hnoohom, N.; Rojroongwasinkul, N.; Kriengsinyos, W. Feasibility Study of an Automated Carbohydrate Estimation System Using Thai Food Images in Comparison with Estimation by Dietitians. *Front. Nutr.* **2021**, *18*, 732449. [CrossRef]

14. Smahel, D.; Elavsky, S.; Machackova, H. Functions of mHealth applications: A user's perspective. *Health Inform. J.* **2019**, *25*, 1065–1075. [CrossRef] [PubMed]

15. Xu, X.; Hou, L.; Guo, Z.; Wang, J.; Li, J. Developing a Chinese food nutrient data analysis system for precise dietary intake management. In Proceedings of the Big Data—BigData 2018: 7th International Congress, Held as Part of the Services Conference Federation, Seattle, WA, USA, 25–30 June 2018.

16. Hamdollahi Oskouei, S.; Hashemzadeh, M. FoodRecNet: A comprehensively personalized food recommender system using deep neural networks. *Knowl. Inf. Syst.* **2023**, *65*, 3753–3775. [CrossRef]

17. Ismail, R.; Yuan, Z. Food Ingredients Recognition through Multi-label Learning. In *Embedded Artificial Intelligence*, 1st ed.; River Publishers: London, UK, 2023; pp. 105–116.

18. Mansouri, M.; Benabdellah Chaouni, S.; Jai Andaloussi, S.; Ouchetto, O. Deep Learning for Food Image Recognition and Nutrition Analysis Towards Chronic Diseases Monitoring: A Systematic Review. *SN Comput. Sci.* **2023**, *4*, 513–530. [CrossRef]

19. Kaur, R.; Kumar, R.; Gupta, M. Food Image-based diet recommendation framework to overcome PCOS problem in women using deep convolutional neural network. *Comput. Electr. Eng.* **2022**, *103*, 108298. [CrossRef]

20. Sarda, E.; Deshmukh, P.; Bhole, S.; Jadhav, S. Estimating food nutrients using region-based convolutional neural network. In Proceedings of the International Conference on Computational Intelligence and Data Engineering: ICCIDE 2020, Hyderabad, India, 12–13 June 2020.

21. Grassi, S.; Alamprese, C. Advances in NIR spectroscopy applied to process analytical technology in food industries. *Curr. Opin. Food Sci.* **2018**, *2*, 17–21. [CrossRef]

22. Cen, H.; He, Y. Theory and application of near infrared reflectance spectroscopy in determination of food quality. *Trends Food Sci. Technol.* **2007**, *18*, 72–83. [CrossRef]

23. del Rio Celestino, M.; Font, R. Using Vis-NIR spectroscopy for predicting quality compounds in foods. *Sensors* **2022**, *22*, 4845. [CrossRef] [PubMed]

24. Lopez, M.G.; García-Gonzalez, A.S.; Franco-Robles, E. Carbohydrate analysis by NIRS-Chemometrics. *Dev. Near-Infrared Spectrosc.* **2017**, *10*, 67208.

25. Qu, J.H.; Liu, D.; Cheng, J.H.; Sun, D.W.; Ma, J.; Pu, H.; Zeng, X.A. Applications of near-infrared spectroscopy in food safety evaluation and control: A review of recent research advances. *Crit. Rev. Food Sci. Nutr.* **2015**, *55*, 1939–1954. [CrossRef]

26. Tellspec Inc. Tellspec-Analysis, Food Safety, Food Database, Food Security. Available online: https://tellspec.com/ (accessed on 13 March 2024).

27. Consumer Physics, Scio. Available online: https://www.consumerphysics.com/ (accessed on 13 March 2024).

28. Zhou, L.; Tan, L.; Zhang, C.; Zhao, N.; He, Y.; Qiu, Z. A portable NIR-system for mixture powdery food analysis using deep learning. *LWT* **2022**, *153*, 112456. [CrossRef]

29. Folli, G.S.; Santos, L.P.; Santos, F.D.; Cunha, P.H.; Schaffel, I.F.; Borghi, F.T.; Filgueiras, P.R. Food analysis by portable NIR spectrometer. *Food Chem. Adv.* **2022**, *1*, 100074. [CrossRef]

30. Hu, H.; Zhang, Q.; Chen, Y. NIRSCAM: A Mobile Near-Infrared Sensing System for Food Calorie Estimation. *IEEE Internet Things J.* **2022**, *9*, 18934–18945. [CrossRef]

31. Ahn, D.; Choi, J.; Kim, H.; Cho, J.; Moon, K.; Park, T. Estimating the composition of food nutrients from hyperspectral signals based on deep neural networks. *Sensors* **2019**, *19*, 1560. [CrossRef] [PubMed]

32. Huang, Q.; Yang, Z.; Zhang, Q. Smart-U: Smart Utensils Know What You Eat. In Proceedings of the IEEE INFOCOM 2018—IEEE Conference on Computer Communications, Honolulu, HI, USA, 16–19 April 2018.

33. Kendall, A.; Gal, Y. What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision? In Proceedings of the Advances in Neural Information Processing Systems, NIPS 2017, Long Beach, CA, USA, 4–9 December 2017.

34. Vartanian, L.R. Impression management and food intake. Current directions in research. *Appetite* **2015**, *86*, 74–80. [CrossRef] [PubMed]

35. Phanich, M.; Pholkul, P.; Phimoltares, S. Food Recommendation System Using Clustering Analysis for Diabetic Patients. In Proceedings of the 2010 International Conference on Information Science and Applications, Seoul, Republic of Korea, 21–23 April 2010.

36. Guo, W.; Wang, J.; Wang, S. Deep multimodal representation learning: A survey. *IEEE Access* **2019**, *7*, 63373–63394. [CrossRef]

37. Chen, W.; Wang, W.; Liu, L.; Lew, M.S. New ideas and trends in deep multimodal content understanding: A review. *Neurocomputing* **2021**, *426*, 195–215. [CrossRef]

38. Xu, M.; Watanachaturaporn, P.; Varshney, P.K.; Arora, M.K. Decision tree regression for soft classification of remote sensing data. *Remote Sens. Environ.* **2005**, *97*, 322–336. [CrossRef]

39. Zhang, S.; Li, X.; Zong, M.; Zhu, X.; Cheng, D. Learning k for knn classification. *ACM Trans. Intell. Syst. Technol.* **2017**, *8*, 1–19. [CrossRef]

40. Seber, G.A.F.; Lee, A.J. *Linear Regression Analysis*; John Wiley & Sons: Hoboken, NJ, USA, 2003.

41. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016.

42. Liaw, A.; Wiener, M. Classification and regression by random Forest. *R News* **2002**, *2*, 18–22.

43. Lee, K.S. Multispectral food classification and caloric estimation using convolutional neural networks. *Foods* **2023**, *12*, 3212. [CrossRef]

44. Han, Y.; Cheng, Q.; Wu, W.; Huang, Z. Dpf-nutrition: Food nutrition estimation via depth prediction and fusion. *Foods* **2023**, *12*, 4293. [CrossRef] [PubMed]

45. Zhu, H.; Gowen, A.; Feng, H.; Yu, K.; Xu, J.L. Deep spectral-spatial features of near infrared hyperspectral images for pixel-wise classification of food products. *Sensors* **2020**, *20*, 5322. [CrossRef] [PubMed]