SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V12I2P107

Volume 12 Issue 2, 73-85, February 2025

Original Article

ProFoodNet: Enhancing Protein Prediction from Food Images Using Advanced Machine and Deep Learning Techniques

P. Josephin Shermila[1], Shanthi Thangam Manukumar[2], C. Reeda Lenus[3], E. Anna Devi[4]

[1]*Department of Artificial Intelligence and Data Science, R. M. K. College of Engineering and Technology, Tiruvallur, Tamilnadu, India.*
[2]*Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamilnadu, India.*
[3]*Department of Physics, S.A. Engineering College, Thiruverkadu, Tamilnadu, India.*
[4]*Department of Electronics and Communication Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamilnadu, India.*

[1]*Corresponding Author : josephinshermilaads@rmkcet.ac.in*

Received: 09 December 2024 Revised: 07 January 2025 Accepted: 08 February 2025 Published: 26 February 2025

Abstract *- In this study, we present ProFoodNet, an advanced protein prediction system from food images leveraging both machine learning and a seven-layer deep Convolutional Neural Network (CNN). ProFoodNet aims to address the challenge of accurately estimating protein content in food items, which is crucial for managing dietary intake in individuals with protein-related health conditions. Our approach utilizes the image database IDPE containing 990 images of various food products. First, we employ gradient-based edge detection operators (Prewitt, Sobel, and Kirsch) to extract image features. Two prediction models are then trained and tested using these features: a deep CNN and a linear regression model using a Support Vector Machine (SVM). The deep CNN model outperforms the SVM-based model by achieving the lowest average prediction error (±1.94), according to experimental results. Our findings highlight the potential of advanced machine and deep learning techniques in improving the accuracy of protein prediction from food images, facilitating dietary management and personalized nutrition advice.*

Keywords *- Protein prediction, Support Vector Machine, Sobel operator, Prewitt operator, Kirsch operator, Deep CNN.*

1. Introduction

Food contains carbohydrates, fats, proteins and other nutrients. Amino acids combine to form proteins. Amino acids synthesize components such as peptide hormones, neurotransmitters and other proteins in the human body. Organs in a body get good structure due to the presence of protein in humans. For proper functioning of body organs protein is essential. One of the challenges nowadays is predicting protein content from food images. Nutrition-related information is required for people with protein-related diseases like marasmus, kwashiorkor, and Maple Syrup Urine Disease. A few diseases such as Chronic Kidney Disease (CKD), Liver Disease, Phenylketonuria (PKU), Gout, Protein-Energy Malnutrition (PEM), Celiac Disease, Diabetes, Heart Disease, Cancer, HIV/AIDS also need the measurement of protein in food intake. Children with Maple Syrup Urine Disease especially require continuous monitoring of protein intake in order to be healthy. Laboratory methods like Sequence-Based Prediction Methods, Structure-Based Prediction Methods and Function Prediction Methods are available. However, a new method is necessary to find the protein in food intake.

This is the rationale behind their direct protein prediction based on food photos. Using the Image Database for Protein Estimation (IDPE), protein may be predicted from food images. Here, a few machine learning methods and deep learning methods are proposed in this research work, namely ProFoodNet: Advanced Protein Prediction from Food Images Using Machine Learning and a 7-Layer Deep Convolutional Neural Network.

2. Literature Survey

Patients in hospitals need to know their nutrition. Protein intake needs to be calculated, especially for the above-mentioned diseases. In earlier days, the patient's attendee calculated this and noted it on a weekly or daily basis. The following section discusses the traditional ways and changes that happened in calculating protein.

One easy technique is to keep track of the amount and kind of food consumed. In the early 20th century, nutrition was manually forecasted based on the type and quantity of food consumed. Because of reporting mistakes by age group, there is an issue with either overestimating or underestimating the amount of food [1]. In order to lessen the issue of reporting errors, a food intake log can be kept. For this reason, a diary is employed to document [2]. The problem of recording in a diary is that it cannot be carried everywhere with the person. The person may not always carry the diary, which causes a delay in recording. That is, on a weekly or monthly basis. This delay, in turn, leads to some reporting errors. Therefore, Bruke et al. [2] suggested an electronic food diary named personal digital assistant. It is advantageous due to its electronic nature of reporting. One of the disadvantages noted here is training the user to use the device. Another disadvantage is improper updating of food intake. Like the diary and pencil method, the personal digital assistant shows the reporting problem. The accuracy of estimating the amount of food consumed by an individual is not guaranteed in either the diary and pencil method or the personal digital assistant.

The accurate measurement of calories and nutrition is required. It is necessary to eradicate the issue of reporting in this century in order to adhere to a therapeutic diet. Researchers have expressed interest in employing image-processing techniques to eradicate reporting inaccuracies. Many researchers have worked to estimate food's nutrition and calorie value from its images.

In [3], the quantity of nutrients and calories is predicted by monitoring an individual's food intake. Initially, the photograph of the food is captured. A checkerboard calibration card is also present along with the food image while capturing the food image. Food volume can be adjusted with the help of the checkerboard image. Food is manually identified, and an image processing method determines the portion size. Using a lookup table, food's calorie and nutrient content are anticipated. The prediction must be made user-friendly by eliminating the need for a checkerboard to estimate the food's volume and the manual intervention needed to determine the type of food.

Gao and Tan [4] unveiled the health-conscious smartphone app. A mobile phone is used to take the image of the food. The type of captured food image is determined through comparison with a food image database. The recognized food's calorie and nutritional information is anticipated using a lookup table. Only the food type's typical nutritional content can be provided by this kind of system.

Marriapan et al. [5] introduced a classification algorithm to identify the type of food item. Food images are segmented. After the image has been segmented, features are extracted. Texture, color, and intensity are the attributes that were retrieved. To forecast the type of food based on the learned attributes, the Support Vector Machine (SVM) classifier is employed. Only 57.55% of the forecasts stated in [5] were correct. The amount of food and the nutrients are anticipated, much like Sun et al. [3].

Aizawa et al. estimate the dietary balance in images of food [6]. A Japanese food-balance chart is used to compare it [7]. Similar food-balance parameters are found in the United States Department of Agriculture (USDA) My Pyramid [8]. This entails photographing the dish being consumed. The recorded images are classified. The pictures are divided into two categories: images of food and non-food images. The Naivie classifier is used to further classify images that come under the food imagine category into categories such as grains, meat, fish, beans, vegetables, and fruit and dairy products. Utilizing the food-balance guide, the nutritional intake for the categorized food image is provided.

An electronic gadget, specifically a camera worn on the breast, is used to forecast the amount of food that the person would ingest [9]. This technological device is worn by the individual. The person using the device must turn it on in order to take an image of the food. The electronic device calculates the amount of food based on the captured image using software. According to Jia et al. [9], the mean relative error between the estimated and real food amount is 2.8% for a total of 100 samples. The main problem with this method is that the person must always wear the e-device, and the device needs to be protected from atmospheric conditions such as rain.

Pouladzadeh et al. calculated the calorie from food images. Initially, [10] identification of food type is done. For identification purposes, the size, color, texture, and shape features are obtained from food images. Using these features with the SVM classifier, the food type is identified. Using the food image, the area of the selected food is computed, followed by its volume and mass, and finally, the calorie is determined using a lookup table. In [3, 5], a checkerboard is used for the volume calculation, but Pouladzadeh et al. used the image of the thumb for the volume prediction. For the volume calculation, using the thumb, the top and side views of the food, along with the thumb, is necessary. This estimation process is complicated to follow in mixed food items. Hence, ingredients in food detection need to be done. Ingredients detection from food images is a boom in 2016 [11].

Pouladzadeh et al. have presented a deep learning neural network for measuring calories [12]. Using a neural network in conjunction with segmentation techniques, including texture, graph cut, and color, improved the accuracy of food recognition for a single food portion [12]. In this case, they used references to estimate the area and volume of the food portion. The area of the food portion is estimated by measuring the distance from which the food is captured, and the amount of the food portion is estimated by measuring the

finger. According to [12], estimating calories using convolutional algorithms takes a long time. It has been found that the processing time is reduced by using deep learning. Anthimopoulos et al. [13] took a step in estimating one of the macronutrient carbohydrates from food images. Along with a reference card, his team took pictures of the food for this study. The volume of the food is calculated using the reference card image. A 3D reconstruction of a food image is used to determine its volume.

Using radial support vector machines, color characteristics, and linear binary pattern features, food kind is determined. The USDA nutritional database is used for carbohydrate estimation. Carbohydrate estimation from food images is remodeled in [14] using a two-view 3D reconstruction method.

Hyperspectral images of the wheat kernel are taken as input in [15] for the purpose of protein measurement. A partial least square regression model is followed for the prediction. Volume prediction and use of a lookup table are eliminated in [15] while predicting the protein from its images. Despite appearing to be comparable to the research done in [15], this publication differs in a few ways, as mentioned below: i) Hyperspectral and photographic pictures are employed in opposition to each other, and ii) the complete food is examined in this study in opposition to the single wheat kernel.

In [16], a Raspberry Pi camera, which is attached to a smart fork, is used by the authors to collect data on food images. Using Bluetooth technology, the data is transmitted to a food database; hence, the calories, starch, fat, and protein are obtained. All the calculations for the above-mentioned nutrition are done using a smartphone, as mentioned in [16]. This is one of the methods for measuring calories from food images for the well-being of humans. As calories vary depending on the ingredients in the food, Turmchokkasam and Chamnongthai identified the food ingredients and obtained their nutritional information, brightness, and thermal conditions from a database. The segmented food is classified using fuzzy logic methods. Based on the area, the total calorie information is obtained as shown in [17].

The creation and use of the AlphaFold model are described in Jumper et al.'s very accurate protein structure prediction with AlphaFold. Key components include training deep neural networks on vast amounts of protein structure data and utilizing attention mechanisms and evolutionary information to capture spatial and sequential dependencies. The model employs an iterative process to refine its predictions, integrating multiple data sources and constraints to enhance accuracy. Advanced optimization techniques and high-performance computing resources were crucial for training the model. AlphaFold system achieved exceptional accuracy in protein structure prediction, greatly outperforming earlier techniques.

Highly precise protein structure prediction using AlphaFold by Jumper et al. details the development and application of the AlphaFold model. Key components include training deep neural networks on vast amounts of protein structure data and utilizing attention mechanisms and evolutionary information to capture spatial and sequential dependencies [18]. The model employs an iterative process to refine its predictions, integrating multiple data sources and constraints to enhance accuracy. Advanced optimization techniques and high-performance computing resources were crucial for training the model. Its results in the Critical Assessment of Protein Structure Prediction (CASP) competition confirmed that the resulting AlphaFold system achieved exceptional accuracy in protein structure prediction, greatly outperforming earlier techniques.

Mohammed AlQuraishi is interested in the Prediction of Protein Structures. It [19] includes handling and comprehending complicated biological data, and it is essential to integrate deep learning models, especially recurrent and convolutional neural networks. The models may discover complex patterns and relationships in the data because they are trained on large databases of known protein structures. In order to anticipate protein folding and interactions, the implementation additionally uses techniques for feature extraction and sequence alignment. The effectiveness of these models is validated through rigorous testing and comparison with traditional prediction methods, showcasing significant improvements in prediction accuracy and efficiency.

The paper Protein-protein interaction prediction with deep learning: A thorough review by Soleymani et al. describes the use of deep learning techniques to predict Protein-Protein Interactions (PPIs). Neural network topologies specifically designed to capture the intricate properties of protein structures and sequences were one of the main focus areas. The research discusses several deep learning models, such as RNN and CNN, which were trained on large PPI datasets [20]. Feature extraction methods were employed to convert protein data into formats suitable for machine learning. The implementation also involved using validation techniques to assess the predictive accuracy of the models. The review highlights significant advancements in the field, emphasizing improved PPI prediction accuracy through deep learning approaches.

The study on food value estimation from images involved several key steps. A comprehensive taxonomy was initially developed to categorize food items based on their nutritional content. Next, a dataset comprising thousands of labelled food images was curated from various sources [21]. The models were trained and fine-tuned using the curated dataset to accurately estimate the nutritional values of different food items. Finally, the system was tested for accuracy and reliability, demonstrating its capability to provide precise food value estimations from image inputs, thus aiding in dietary

assessment and management. The creation of a unique approach for calculating food nutrition is described in [22] DPF-Nutrition: Food Nutrition Estimation by Depth Prediction and Fusion by Han et al. The implementation involved using depth prediction techniques to capture 3D information on food items, improving accuracy in portion size estimation. The system integrates multiple data sources, including RGB images and depth maps, using a fusion approach to enhance the precision of nutritional estimates.

Machine learning algorithms were trained on extensive datasets to recognize different food types and their nutritional values. The model was validated through rigorous testing, demonstrating its effectiveness in providing reliable nutrition information, aiding dietary assessment and management.

Hurry Sign and Noman's Deep Neural Network-Based Food Protein Subcellular Prediction and IoT-Based Data Gathering Ali investigates the prediction of food proteins' subcellular location using Deep Neural Networks (DNNs). Knowing how proteins work and how they contribute to different biological processes depends on this prediction [23].

The authors employ an Internet of Things (IoT)-based data collection system to gather extensive and diverse datasets, including protein sequences, cellular localization signals, and environmental factors. These data are pre-processed and fed into deep neural networks specifically designed to handle large, complex datasets.

The DNN architecture includes multiple layers of convolutional and recurrent neural networks, adept at capturing intricate patterns and dependencies within the data. The model is trained and validated using standard performance metrics like accuracy, precision, and recall.

The results indicate that the DNN approach, combined with IoT-based data collection, significantly improves the accuracy of subcellular localization predictions, offering valuable insights for food protein research and potential applications in biotechnology and medicine.

The article Recent Progress of Protein Tertiary Structure Prediction by QiqigeWuyun et al. delves into the latest advancements in predicting protein tertiary structures, a critical aspect of understanding protein function and interactions. The implementation focuses on leveraging cutting-edge computational techniques and machine learning models to enhance prediction accuracy.

The authors integrate various data sources, including amino acid sequences, known structural motifs, and evolutionary information. They use deep learning architectures to handle and evaluate these intricate datasets, including Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs). The models are trained on extensive protein databases, utilizing advanced algorithms like AlphaFold and RoseTTAFold to predict protein folding patterns [24].

Performance is assessed using metrics such as Global Distance Test-Total Score (GDT-TS) and Root-Mean-Square Deviation (RMSD). As evidence of the effectiveness of contemporary machine learning techniques, the results show notable increases in prediction accuracy. This progress in protein tertiary structure prediction holds promise for advancing biological research and drug discovery.

The article Optimizing Food Protein Prediction for Drug Composition Using Feature Fusion Techniques by Malik Arshad and Andrew McCullum focuses on enhancing the accuracy of predicting food protein interactions within drug compositions. The implementation revolves around leveraging advanced feature fusion techniques, which combine multiple data sources to improve predictive performance.

Initially, the authors collect and preprocess diverse datasets, including protein sequences, structural information, and biochemical properties. They employ various feature extraction methods to derive meaningful representations from these datasets. These features include primary sequence patterns, secondary structure elements, and physicochemical properties.

The core of the implementation lies in integrating these heterogeneous features using fusion techniques [25]. The authors experiment with several fusion strategies, such as early fusion, where features are combined at the input level, and late fusion, where predictions from individual models are merged. They also explore hybrid fusion approaches that integrate features at multiple stages of the prediction pipeline. In order to train the prediction models, the authors use deep learning architectures like CNNs and RNNs as well as machine learning methods like SVM and random forests. A carefully selected dataset is used to train and validate the models, guaranteeing their generalizability and robustness.

The F1 score, recall, accuracy, and precision are among the metrics used to evaluate performance. The findings show that feature fusion techniques greatly improve prediction accuracy in contrast to conventional single-source feature methods. Implementing these techniques shows promise in optimizing drug composition by accurately predicting food-protein interactions, potentially leading to more effective and safer pharmaceutical formulations.

Prediction of protein from food can be investigated through chemical-based evaluation and image processing. Chemical-based methods for the prediction of proteins are a lengthy process. Also, the common person cannot always test in a laboratory. Everyone has a mobile nowadays, so it is easy

to capture the image and find the protein easily. One more reason is that children with protein-related diseases can find the amount of protein they need whenever they wish. Hence, the research on the measurement of protein from food is carried out using an image processing technique using MATLAB.

The document's remaining sections are arranged as follows: The proposed framework is presented in Section 3, the protein prediction results are displayed in Section 4, and the study's conclusions are compiled in Section 5.

3. Proposed Method

3.1. Image Database

Proteins in food photos are predicted using the Image Database for Protein Estimation (IDPE). The 990 images in the database are made up of nine different foods. Food images comprise images of milk powders and images of health drink powders. The IDPE database consists of images of food items such as Boost, Bournvita, Complan, Dexolac, Farex, Horlicks Lactogen, Nan and Similac. Figure 1 shows a few images of the IDPE database. IDPE database can also predict calories, macronutrients, and micronutrients.

IDPE database images are in JPG format. The size of the images is 3120 x 4160. In addition, Table 1 shows the food items in the IDPE database along with their names, properties and the amount of protein in 100 grams.

Fig. 1 Sample images of IDPE

Table 1. Food items and the protein content

Name of the food item	Fine/Corse	Protein content per 100 grams
Lactogen	Fine	15.20
Nan	Fine	15.30
Similac	Very Fine	15.50
Farex	Fine	17.00
Dexolac	Very Fine	12.00
Horlicks	Coarse	11.00
Boost	Coarse	06.50
Bournvita	Coarse	06.00
Complan	Coarse	19.00

3.2. ProFoodNet: Advanced Protein Prediction from Food Images

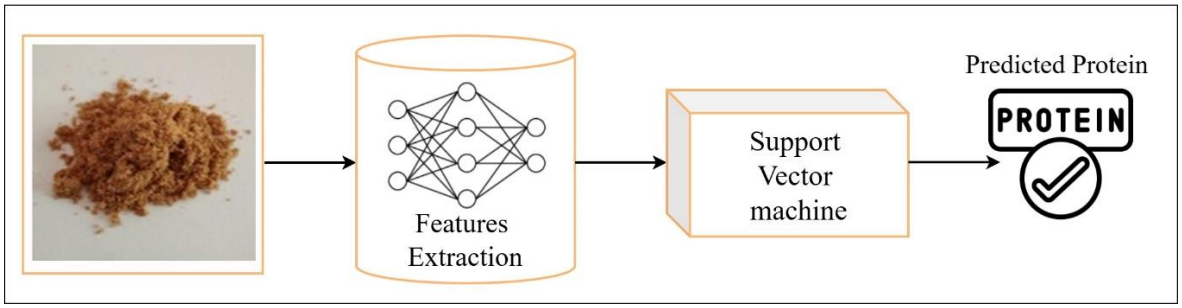

Fig. 2 Proposed machine learning model

Real-world applications of the technology include supporting individuals with protein-specific dietary demands, such as athletes enhancing performance, patients managing chronic kidney disease, or those requiring exact protein intake for malnutrition recovery. ProFoodNet might also be shown in contexts such as hospitals, fitness facilities, or mobile health applications, showcasing how it integrates smoothly into everyday food planning and monitoring**.** Both machine learning and deep learning models are used in this study to predict proteins. The machine learning model for the suggested task is displayed in Figure 2.

Figure 2 shows the Deep learning model, namely ProFoodNet: Advanced Protein Prediction from Food Images used for the protein prediction. This method is briefed in section 3.4.

3.3. Feature Extraction

There are differences between the types of information that are collected from photos and those that are perceived by human eyes. If the objects and features in an image cannot be precisely compared by placing them next to each other, humans are particularly bad at determining their color or

brightness. The way that human vision responds to the relative size, angle, or location of many things is essentially comparative rather than quantitative; it cannot provide numerical measurements unless one of the reference items is a measuring scale. The characteristics of an image are represented by some information called the image features. Feature play an important role in the area of image processing. Color, shape, texture, contrast and size are some examples of features. Texture features can be extracted to analyze the different structures in an image. Image brightness changes or discontinuity can be sharply identified using edge characteristics. Edge features can be derived from images to model some constructions.

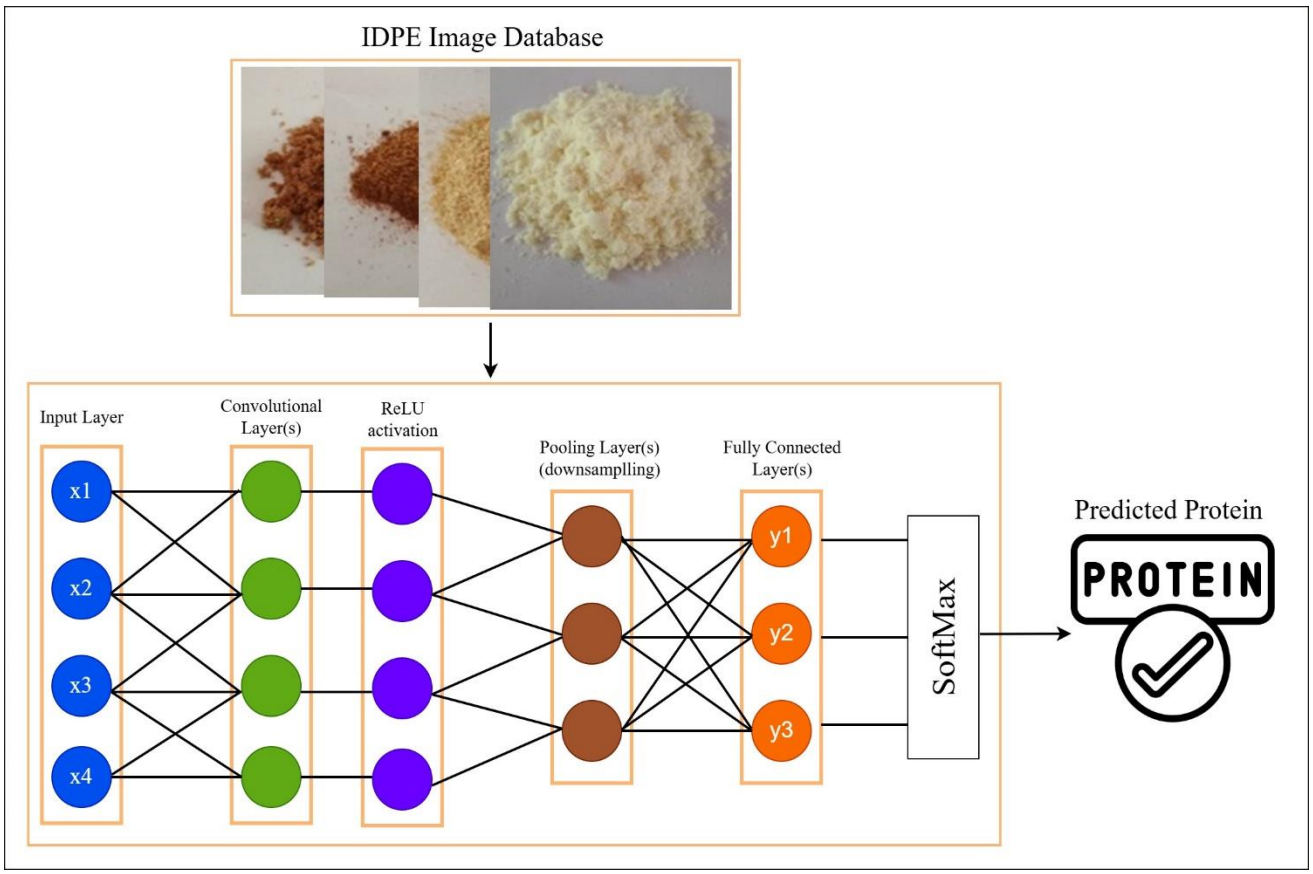

Fig. 3 ProFoodNet: Advanced protein prediction from food images

Prewitt mask is used to detect the horizontal and vertical edges. They measure the 2D spatial gradient values. Sobel mask, similar to Prewitt, has more weight and hence detects more edges. Additionally, the intensity values' first-order derivative is produced. The edges of the Kirsch mask can be found in eight distinct directions: north, west, east, south, northeast, northwest, southeast, and southwest. To accomplish this, turn the mask 45 degrees.

3.3.1. Gradient Magnitude and Direction Using Prewitt Operator

The usage of gradient values is one of the characteristics taken into account for protein prediction. The 2D spatial gradient values are detected using the Prewitt mask, which is employed for this purpose. Using the Prewitt operator, one may determine the direction of the greatest potential shift in intensity. The Prewitt mask can be used to analyze abrupt or gradual feature changes at a spot. Two templates are available for the Prewitt edge detection operator. Both the templates are of size (3x3). The vertical template along the three rows and the horizontal template along the 3 columns are given by [26] Prewitt & Mendelsohn (1966).

The template and the image are convolved to generate the 2D spatial derivatives. Consider an image (s) and its gradient (∇s). Additionally, let GDx and GDy represent the picture's horizontal and vertical derivatives, respectively. The following is the location of the derivatives GDx and GDy.

$$\nabla s = \begin{bmatrix} GD_x \\ GD_y \end{bmatrix} = \begin{bmatrix} \frac{\partial s}{\partial x} \\ \frac{\partial s}{\partial y} \end{bmatrix} \quad (1)$$

The gradient magnitude and gradient direction $z(x, y)$ are obtained as follows.

$$mag(\nabla s)=\left[GD_x^2 GD_y^2\right]^{1/2} \quad (2)$$

$$z(x,y) = tan^{-1}\left(GD_y/GD_x\right) \quad (3)$$

Prewitt kernels compute the gradient. It is the product of averaging and then differentiation. The final resultant has a smoothening effect. Using equations 1, 2, and 3, the statistical measurements examined in this study are correlation, mean, and standard deviation.

3.3.2. Gradient Magnitude and Direction Using Sobel Operator

It calculates an approximate gradient for the picture intensity function. The Sobel operator is a discrete differentiation operator. In the classical Sobel operator situation, each image has a template, GDx and GDy. The gradient estimate in the x-direction is represented by GDx, and the gradient estimation in the y-direction by GDy, as per equation 1. Equation 2 can be used to get the absolute gradient. Equation 3 can be utilized to determine the gradient's direction. The mask used for the Sobel operator is shown below.

Due to its ability to detect edges in images containing redundant information, such as noise, the Sobel operator is the gradient-based operator most commonly used for edge detection. Because each image is differentially separated by two rows and columns, the edge components on both sides are enhanced, giving the edges a thick and vivid appearance. The gradient direction is obtained using equations 1, 2, and 3.

3.3.3. Gradient Magnitude and Direction Using Kirch Operator

Krisch mask is also a derivative mask. It takes a single mask and rotates in 8 different directions. The greatest of the eight orientations determines the image edge magnitude. Applying the mask involves angles of 0, 45, 90, 135°, 180°, 225°, 270 degrees, 315°, and 360°. This technique offers important edge properties when used for database photos.

The numerous compass directions can be obtained at 45-degree increments. The compass indicates the following directions: North, South, East, West, North East, North West, South East, and South West. Prewitt, Sobel, and Kirsch are used to extract statistical features from the RGB, red, green, blue, and gray images of the IDPE.

3.4. Protein Estimation

From IDPE food images, the protein content is estimated using two different methods. This is the first protein prediction method that uses the suggested linear regression with support vector machine technique. Using deep learning as an advanced technique for this study, ProFoodNet is the second approach used for protein prediction: Protein Prediction with Advanced Images of Food.

In the first method, the steps followed are 1) IDPE images are resized, 2) a mask is applied to the images, and the gradient features are obtained. 3)statistical features are obtained from the gradient values. 4) A model of linear regression using support vector machines. Step four is utilized to determine the protein content and make predictions based on the properties of the image. Figure 2 outlines the procedures for protein prediction using support vector machines and linear regression.

Protein content serves as the independent variable in this study, while the image's statistical properties serve as the dependent variable. The model learns about the correlation between protein content and statistical image attributes throughout the training process.

A ten-fold cross-validation scheme is also used in this investigation. Hence, training to test ratio is 90:10. Ten sets of statistical features are separated here, and they are mutually exclusive. Nine sets are utilized for testing during the investigation, whereas one set is used for testing. The regression model is trained and tested 10 times in such a way that each set acts as a test at any one time. The averaged value of ten-fold cross-validation produces the estimation.

Prewitt, Sobel, and Kirsch convolution are used independently for each image component in our experiment, yielding gradient features for each. Statistical features such as 2-dimensional mean, standard deviation and autocorrelation are derived for a better protein prediction.

In the second method, the steps are 1) IDPE images are resized, 2) Images are fed as input to deep CNN. Figure 3 shows the protein prediction using deep CNN. In this investigation, deep CNN with 7 layers is followed. The layers are 1) the Input layer, 2) the Convolutional layer, 3) the Rectification layer, 4) the Pooling Layer, 5) the Fully connected layer, and 6) the Regression layer.

The IDPE images are resized to 256x256 and are given as input to the input layer of deep CNN. Then, the images are convolved using a 3x3 mask, followed by the rectification process. In the rectification layer, the activation process is done as a threshold process. Next is the average pooling process, in which the averaging is done using a mask size 2x2 and with stride 2. In this investigation using deep CNN,40 epochs having 200 iterations and with a learning rate of 0.001is followed. For the purpose of training the model, the steepest gradient descent algorithm is used.

4. Results and Discussion

The first approach yielded average protein prediction errors using SVM training and testing of linear regression. Table 2 shows the statistical features and the abbreviations used in this paper.

Tables 3, 4, and 5 display the average prediction errors derived from training and testing linear regression using each combination of image attributes of the four image components of Prewitt, Sobel, and Kirsch convolution, respectively. Table 3 displays the average protein prediction error derived from combining statistical measurements of gradient magnitude and gradient direction based on the Prewitt operator. In this case, the prediction error is less when the combination of MGM, SGM, MGD and SGD is used for predicting the protein. The findings indicate that proteins can be predicted with a minimum average error of ±2.68 based on the gradient magnitude and gradient direction of the red color component of the database image.

Table 2. Statistics on the direction and degree of gradients

Statistical Feature	Abbreviation
Mean of Gradient Magnitude	MGM
Standard Deviation of Gradient Magnitude	SGM
Autocorrelation of Gradient Magnitude	CGM
Mean of Gradient Direction	MGD
Standard Deviation of Gradient Direction	SGD
Autocorrelation of Gradient Direction	CGD

Table 3. Prewitt feature-derived average prediction error

Statistical Measures of Gradient Magnitude and Gradient Direction Feature	Average Prediction Error			
	Image Component			
	Gray Color Component	Red Color Component	Green Color Component	Blue Color Component
MGM	±2.98	±2.97	±2.95	±2.95
SGM	±2.92	±2.90	±2.88	±2.88
CGM	±2.81	±2.80	±2.84	±2.84
MGD	±2.77	±2.82	±**2.76**	±**2.76**
SGD	±2.79	±2.78	±2.77	±2.77
CGD	±2.81	±2.89	±2.89	±2.90
MGM, MGD	±2.99	±2.94	±2.93	±2.93
SGM, SGD	±2.81	±2.72	±2.79	±2.79
CGM, CGD	±**2.75**	±2.82	±2.84	±2.84
MGM, SGM, CGM	±3.06	±3.06	±3.01	±3.01
MGD, SGD, CGD	±2.81	±2.82	±2.81	±2.81
MGM, SGM, MGD, SGD	±3.00	±**2.68**	±2.78	±2.78
MGM, SGM, CGM, MGD, SGD, CGD	±3.03	±2.90	±2.98	±2.98

Table 4. Mean prediction error as determined by the sobel feature

Statistical Measures of Gradient Magnitude and Gradient Direction Feature Tests	Average prediction error			
	Image component			
	Gray Color Component	Red Color Component	Green Color Component	Blue Color Component
MGM	±2.97	±3.01	±2.97	±2.95
SGM	±2.93	±2.91	±2.90	±2.87
CGM	±2.77	±2.84	±2.81	±2.87
MGD	±2.80	±**2.70**	±**2.75**	±**2.77**
SGD	±2.76	±2.76	±2.75	±2.80
CGD	±**2.74**	±2.79	±2.81	±2.84
MGM, MGD	±2.98	±2.94	±3.01	±2.96
SGM, SGD	±2.87	±2.83	±2.87	±2.85
CGM, CGD	±2.74	±2.82	±2.84	±2.81
MGM, SGM, CGM	±2.97	±3.05	±3.01	±2.95
MGD, SGD, CGD	±2.80	±2.82	±2.80	±2.86
MGM, SGM, MGD, SGD	±2.88	±2.95	±2.94	±2.94
MGM, SGM, CGM, MGD, SGD, CGD	±2.90	±3.00	±2.90	±2.93

The average protein prediction error obtained using the combinations of statistical measures of gradient magnitude and gradient direction based on the Sobel operator is shown in Table 4. In this case, the prediction error is lower when MGD is used to predict the protein.

Based on the Kirsch operator, Table 5 displays the average protein prediction error derived from the combinations of statistical measures of gradient magnitude and gradient direction. Using CGM to estimate the protein reduces the prediction error in this situation. The findings demonstrate that proteins can be predicted with a minimum average error of ±2.71 based on the gradient magnitude and gradient direction of the red color component of the database image.

Comparing the three results based on the operators for the prediction of protein, the red colour component of the Prewitt template produces the best result with a minimum prediction error of ±2.68.

The outcomes of the second technique are the average protein prediction errors obtained by using deep CNN to train and test the database picture. Protein prediction errors obtained when images are given directly as input to deep CNN with 7 layers are shown in Table 6.

Figures 4, 5 and 6 show the protein prediction error obtained while experimenting with Prewitt, Sobel and Krisch features. Figure 7 shows the Average Protein Prediction Error obtained with the CNN model formed in 7 layers.

Table 5. Kirsch feature-derived average prediction error

Statistical Measures of Gradient Magnitude and Gradient Direction Feature	Average prediction error			
	Image component			
	Gray color component	Red color component	Green color component	Blue color component
MGM	±2.80	±2.81	±2.81	±2.78
SGM	±2.80	±2.80	±2.80	±2.81
CGM	±2.79	±**2.71**	±2.73	±2.79
MGD	±2.81	±2.82	±2.80	±2.79
SGD	±**2.74**	±2.81	±2.77	±2.81
CGD	±2.87	±2.74	±2.76	±2.79
MGM, MGD	±2.80	±2.87	±2.93	±2.96
SGM, SGD	±3.52	±3.56	±3.57	±3.67
CGM, CGD	±2.79	±2.81	±**2.72**	±**2.73**
MGM, SGM, CGM	±3.15	±2.95	±3.18	±3.32
MGD, SGD, CGD	±2.81	±2.82	±2.77	±2.85
MGM, SGM, MGD, SGD	±2.98	±3.42	±3.63	±3.60
MGM, SGM, CGM, MGD, SGD, CGD	±3.53	±3.18	±3.52	±3.64

In this case, the minimum average error of ±1.94 is obtained when the RGB image is given directly as input to deep CNN. Deep CNN with 7 layers produces less error compared to the Prewitt operator. This is 0.74 less than the minimum error obtained with the Prewitt operator. ProFoodNet: Advanced Protein Prediction Method followed here predicts protein with less error.

4.1. Comparison of Different Methodologies

Figure 8 displays the comparative outcomes of protein prediction. The best findings are used for this comparison. The CNN model is found to have the lowest prediction error, at 1.94. Compared to the CNN model, the Prewitt, Sobel, and Krisch models provide 0.74, 0.76, and 0.77 more errors, respectively. Hence, the best model for predicting proteins from images of food is ProFoodNet: Advanced Protein Prediction Method, particularly for IDPE images that include milk and health drink powders.

Table 6. Average protein prediction error obtained by deep CNN with 7 layers

Image Component	Average Prediction Error
	Average Pooling
Gray	±2.05
Red	±2.14
Green	±2.01
Blue	±1.97
RGB	±1.94

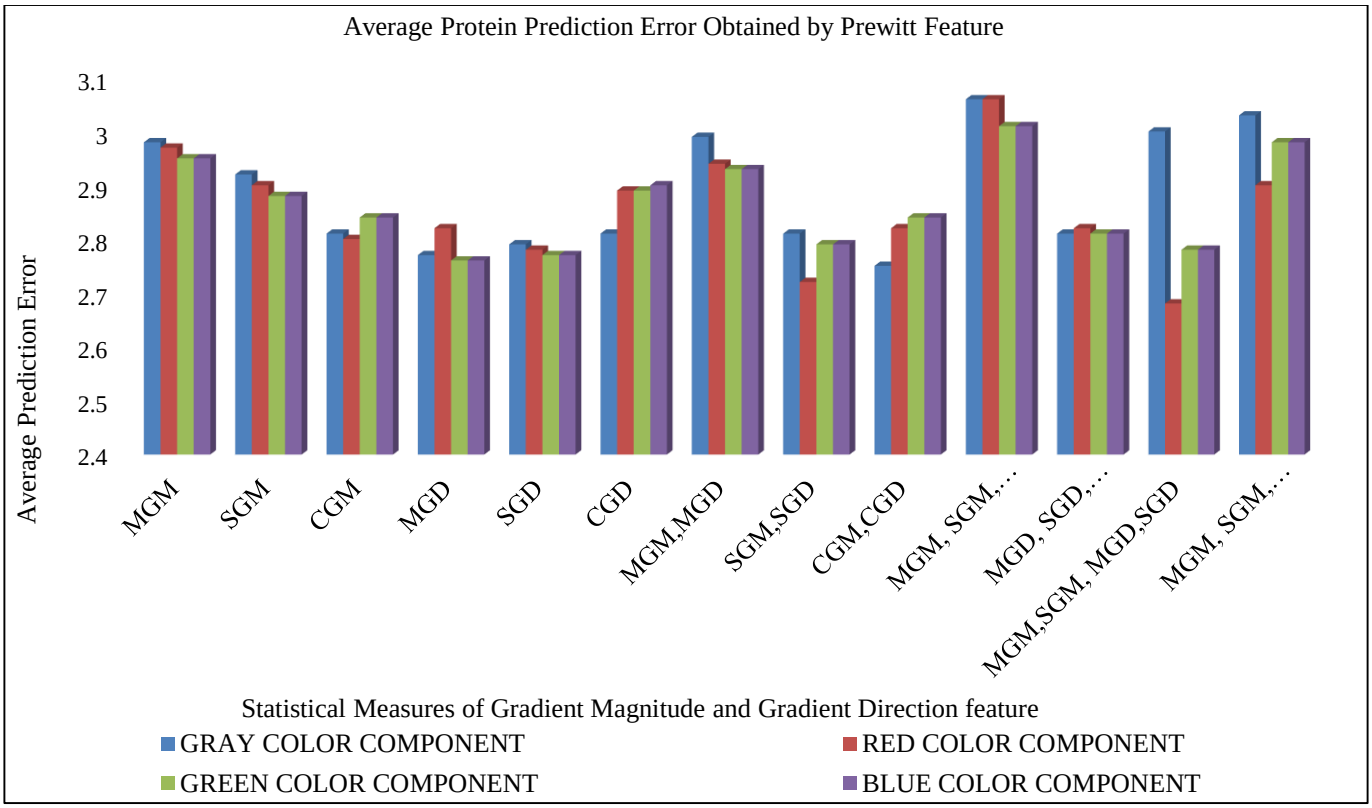

Fig. 4 Results of prewitt feature

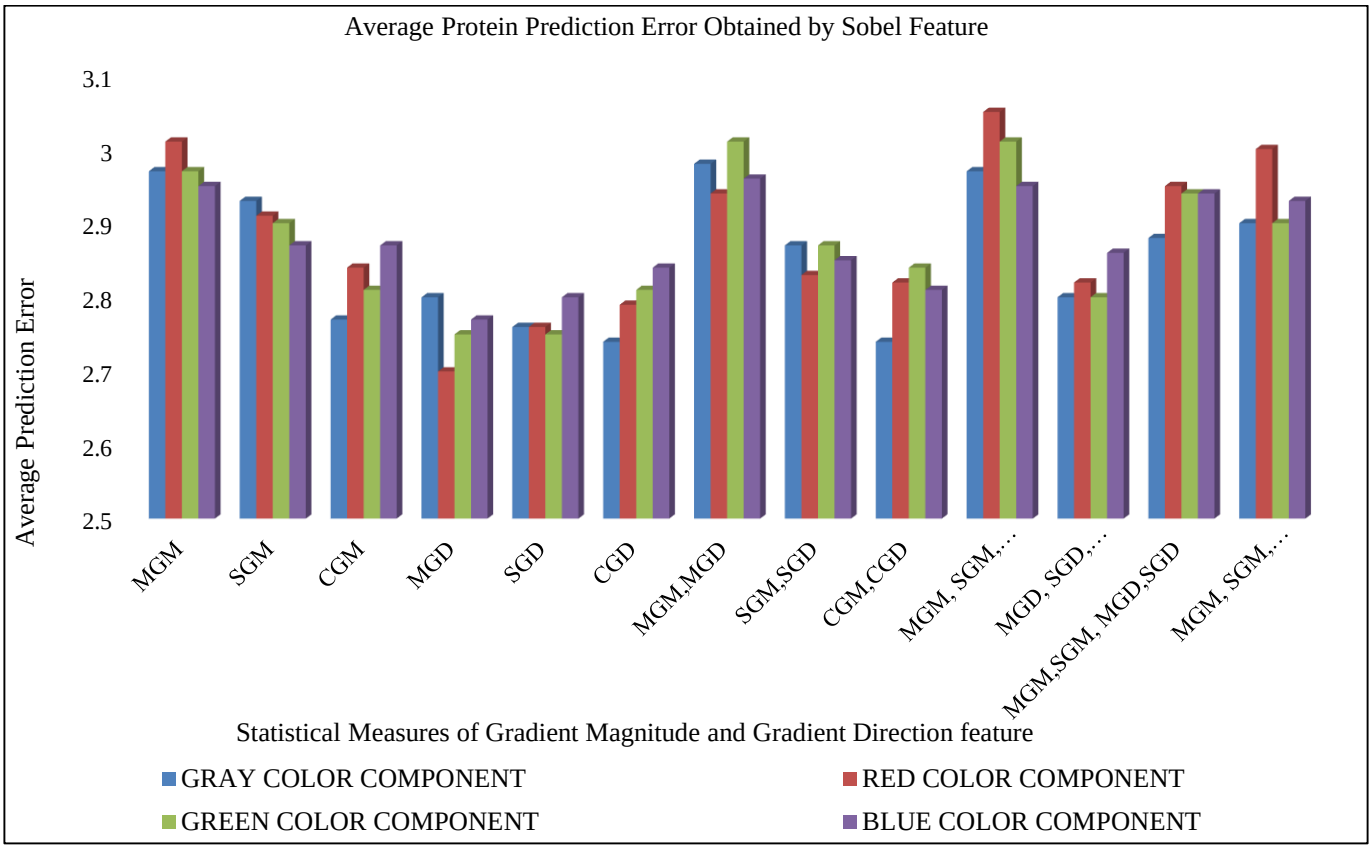

Fig. 5 Results of sobel feature

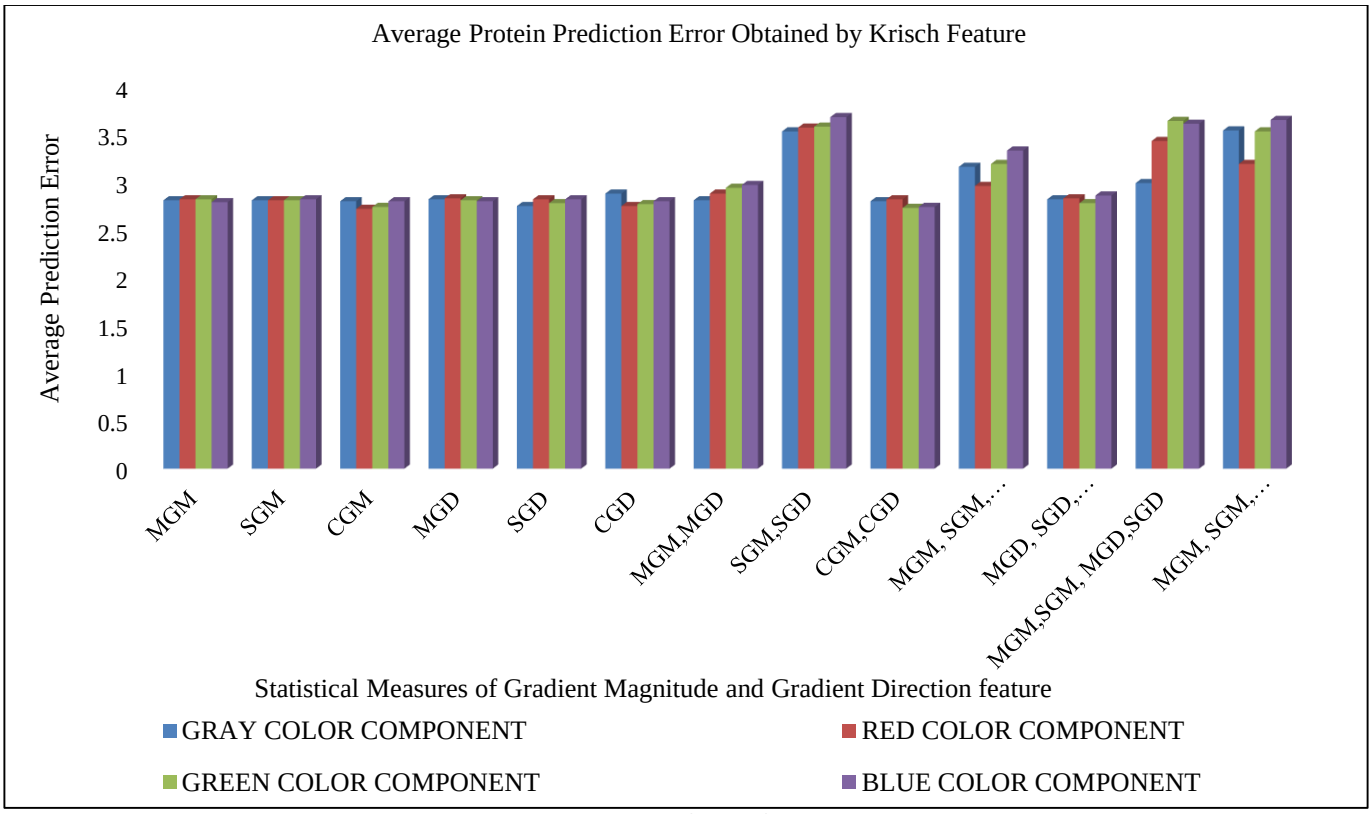

Fig. 6 Results of krisch feature

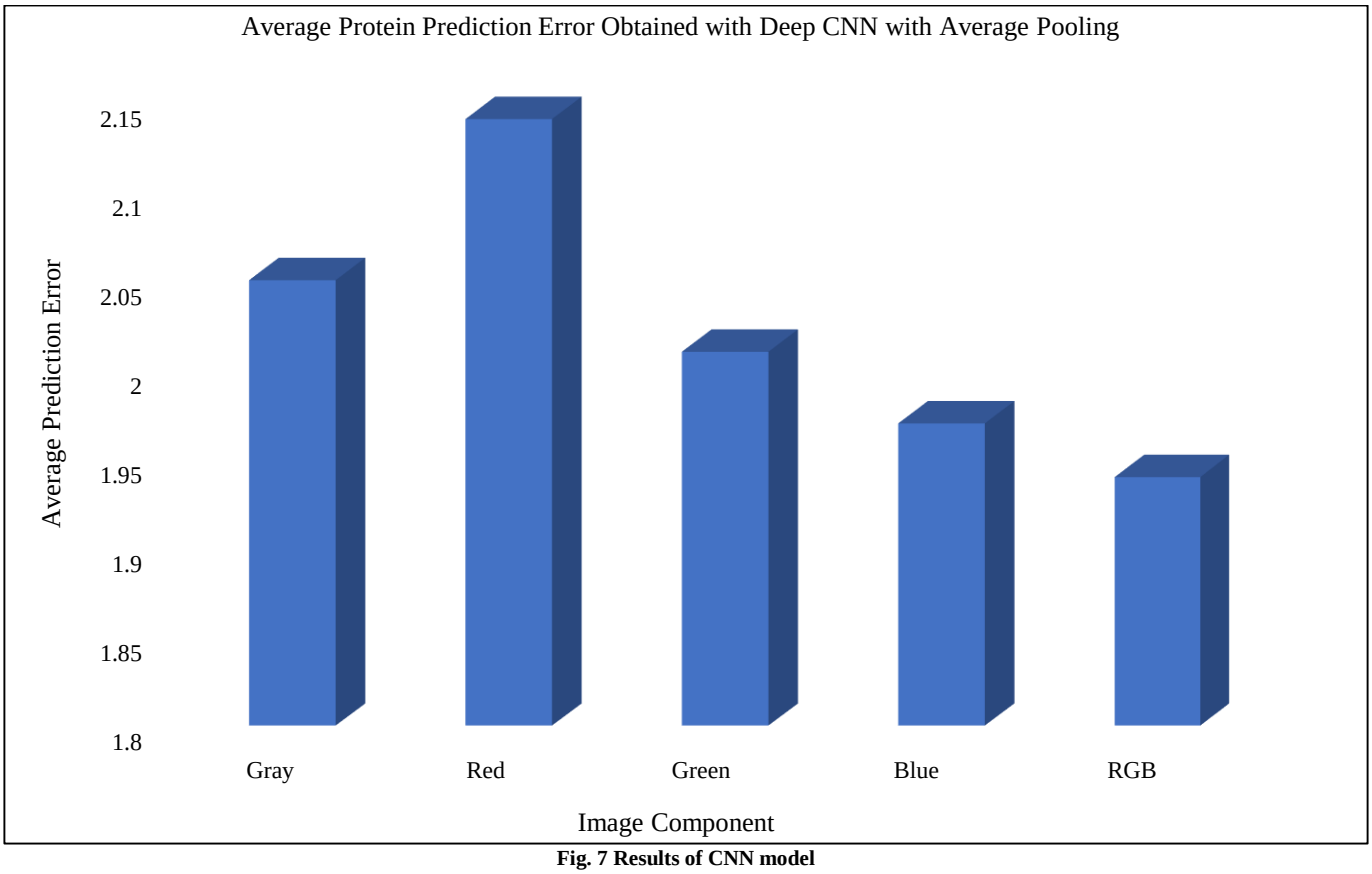

Fig. 7 Results of CNN model

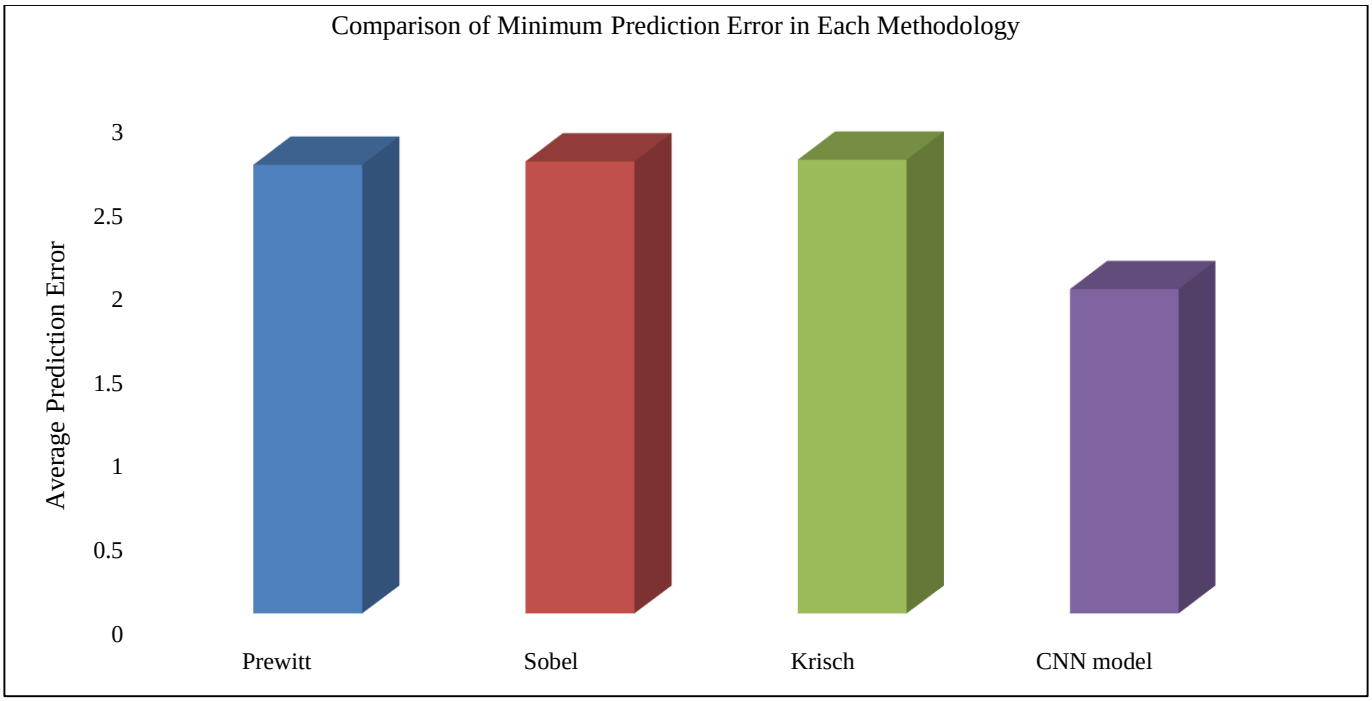

Fig. 8 Comparison of best average prediction error

5. Conclusion and Future Work

The evaluation of protein from food is done in this study using an IDPE database. The protein content of food products is predicted using both deep CNN and a linear regression model using support vector machines. When the linear regression system is modeled using mean and standard deviation (gradient magnitude and direction) characteristics, the Prewitt mask applied to the red color component yields a minimum average protein prediction error of ±2.68By using deep CNN, it is possible to observe a minimum average protein prediction error of ±1.94. Deep learning has a reputation for being able to identify patterns in photos, but it is more effective at predicting the protein content of images when it uses image attributes as input. Additional features, particularly one that conveys the distinctive change information of light when reflected by protein, can be taken to improve the accuracy of protein prediction. Research can be conducted by adding more photos of food from fruits, vegetables, mixed foods, and other sources to the existing study database.

Acknowledgments

The supervisor provided invaluable direction and unwavering support throughout this research, for which the author is deeply grateful.

References

[1] M.B.E. Livingstone, P.J. Robson, and J.M.W. Wallace, "Issues in Dietary Intake Assessment of Children and Adolescents," *British Journal of Nutrition*, vol. 92, no. S2, pp. S213-S222, 2004. [CrossRef] [Google Scholar] [Publisher Link]

[2] Lora E. Burke et al., "Self-Monitoring Dietary Intake: Current and Future Practices," *Journal of Renal Nutrition*, vol. 15, no. 3, pp. 281-290, 2005. [CrossRef] [Google Scholar] [Publisher Link]

[3] Mingui Sun et al., "Determination of Food Portion Size by Image Processing," *30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vancouver, BC, Canada, pp. 871-874, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[4] Chunming Gao, Fanyu Kong, and Jindong Tan, "Healthaware: Tackling Obesity with Health Aware Smart Phone Systems," *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Guilin, China, pp. 1549-1554, 2009. [CrossRef] [Google Scholar] [Publisher Link]

[5] Anand Mariappan et al., "Personal Dietary Assessment using Mobile Devices," *Computational Imaging VII*, vol. 7246, pp. 294-305, 2009. [CrossRef] [Google Scholar] [Publisher Link]

[6] Kiyoharu Aizawa et al., "Food Balance Estimation by using Personal Dietary Tendencies in a Multimedia Food Log," *IEEE Transactions on Multimedia*, vol. 15, no. 8, pp. 2176-2185, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[7] Food Balance Guide, Ministry of Agriculture, Forestry and Fisheries of Japan. [Online]. Available: https://www.maff.go.jp/j/balance_guide/

[8] MyPlate & Food Pyramid Resources, United States Department of Agriculture. [Online]. Available: https://fnic.nal.usda.gov/dietary-guidance/

[9] Wenyan Jia et al., "Accuracy of Food Portion Size Estimation from Digital Pictures Acquired by a Chest-Worn Camera," *Public Health Nutrition*, vol. 17, no. 8, pp. 1671-1681, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[10] Parisa Pouladzadeh, Shervin Shirmohammadi, and Rana Al-Maghrabi, "Measuring Calorie and Nutrition from Food Image," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 8, pp. 1947-1956, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[11] Hongsheng He, Fanyu Kong, and Jindong Tan, "DietCam: Multiview Food Recognition using a Multikernel SVM," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 3, pp. 848-855, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[12] Parisa Pouladzadeh et al., "Food Calorie Measurement using Deep Learning Neural Network," *IEEE International Instrumentation and Measurement Technology Conference Proceedings*, Taipei, Taiwan, pp. 1-6, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[13] Marios Anthimopoulos et al., "Computer Vision-Based Carbohydrate Estimation for Type 1 Patients with Diabetes using Smartphones," *Journal of Diabetes Science and Technology*, vol. 9, no. 3, pp. 507-515, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[14] Joachim Dehais et al., "Two-View 3D Reconstruction for Food Volume Estimation," *IEEE Transactions on Multimedia*, vol. 19, no. 5, pp. 1090-1099, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[15] Nicola Caporaso, Martin B. Whitworth, and Ian D. Fisk, "Protein Content Prediction in Single Wheat Kernels using Hyperspectral Imaging," *Food Chemistry*, vol. 240, pp. 32-42, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[16] K.S. Dheeraj Belliappa et al., "Food Recognition and Analysis using Image Processing," *International Journal of Advance Research, Ideas and Innovations in Technology*, vol. 4, no. 3, pp. 973-978, 2018. [Google Scholar] [Publisher Link]

[17] Sirichai Turmchokkasam, and Kosin Chamnongthai, "The Design and Implementation of an Ingredient-Based Food Calorie Estimation System using Nutrition Knowledge and Fusion of Brightness and Heat Information," *IEEE Access*, vol. 6, pp. 46863-46876, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[18] John Jumper et al., "Highly Accurate Protein Structure Prediction with AlphaFold," *Nature*, vol. 596, pp. 583-589, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[19] Mohammed AlQuraishi, "Machine Learning in Protein Structure Prediction," *Current Opinion in Chemical Biology*, vol. 65, pp. 1-8, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[20] Farzan Soleymani et al., "Protein-Protein Interaction Prediction with Deep Learning: A Comprehensive Review," *Computational and Structural Biotechnology Journal*, vol. 20, pp. 5316-5341, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[21] Jamalia Sultana et al., "A Study on Food Value Estimation from Images: Taxonomies, Datasets, and Techniques," *IEEE Access*, vol. 11, pp. 45910-45935, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[22] Yuzhe Han et al., "DPF-Nutrition: Food Nutrition Estimation via Depth Prediction and Fusion," *Foods*, vol. 12, no. 23, pp. 1-18, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[23] Noman Ali Hurry Sign, "Food Protein Subcellular Prediction Using Deep Neural Networks and IoT-Based Data Collection," [Google Scholar]

[24] Qiqige Wuyun et al., "Recent Progress of Protein Tertiary Structure Prediction," *Molecules*, vol. 29, no. 4, pp. 1-28, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[25] Malik Arshad, and McCullum Andrew, "Optimizing Food Protein Prediction for Drug Composition using Feature Fusion Techniques," [Google Scholar]

[26] J. M. Prewitt, and M. L. Mendelsohn, "The Analysis of Cell Images," *Annals of the New York Academy of Sciences*, vol. 128, no. 3, pp. 1035-1053, 1966. [CrossRef] [Google Scholar] [Publisher Link]