FOOD IMAGE ANALYSIS USING DEEP LEARNING

T. Sai Sindhuja[1]
M. Tech
Department of Computer Science and Engineering
Gayatri Vidya Parishad College of Engineering (Autonomous)

Dr. D.N.D. Harini[2]
Associate Professor
Department of Computer Science and Engineering
Gayatri Vidya Parishad College of Engineering (Autonomous)

Dr. G. Satya Keerthi[3]
Associate Professor
Department of Computer Science and Engineering
Gayatri Vidya Parishad College of Engineering (Autonomous)

ABSTRACT

Healthy lifestyle is an important factor for disease free survival in this pandemic era. To maintain proper health, it is essential to consume good nutritious food. This research embarks on a pivotal journey to harness the capabilities of automatic food image recognition systems, with a profound focus on nutrient identification. Such a system stands poised to usher in a transformative wave within the domains of computer vision, dietary analysis, and fitness monitoring. The objective of this study is to train a deep learning model which identifies a food image captured with any camera device and generate a nutrient estimate report. The proposed model is a two-step process, firstly it recognizes a processed food image using a deep learning model, then it generates a dietary assessment report based on a synthesized nutrient value dataset taking USDA national nutrient database as a standard source. Transfer learning is applied on VGG16 model architecture and experimental evaluation is conducted using subset of food image dataset FOOD-101. Remarkable outcomes emerge through the adaptation of a new CNNovaNet architecture, yielding exceptional results, particularly when applied to the complete FOOD-101 dataset.

INDEX TERMS: *Food Images, Convolutional Neural Network, Image Recognition, VGG16, CNNovaNet, Health Monitoring, Nutrition Analysis.*

INTRODUCTION

Lifestyles led by most people today pose several health risks. Due to the changes in nutrition consumption, technological advancements, physical activity, more modern facilities people are facing health related issues right from a very young age compared to our previous generations. Over the years change is inevitable for the betterment of mankind, but at the same time health consciousness should also be improved. The types of foods and packaged items that are consumed nowadays pose great health risks on most individuals. Most diseases that currently affect people are associated with poor nutrition and lifestyle. Some are curable while others can only be controlled. Chronic diseases are illnesses that are very persistent and last for a long period. They are caused by dietary factors like poor nutrition, behavioural factors like smoking and social factors as race, socioeconomic status and education level. These illnesses are a major cause of deaths globally as they constitute over 60% of all deaths annually [13]. In recent times, after the covid-19 pandemic, people have begun to explore ways to cultivate healthier living habits for improved well-being. Apart from various methods to enhance one's quality of life, like regular exercise, adequate hydration, yogic practices, stress management, intermittent fasting, limit screen time, '*Mindfull Eating*' brings a quick drastic change in day-to-day living. This work aims to facilitate best method to make users aware of the food they eat.

According to World Health Organization '*Nutrition*' is defined as a physiological process by which substances in food are transformed into body tissues and provide energy for the full range of physical and mental activities that make up human life. A good nutritious food must include fruits, vegetables, grains, legumes, protein rich substances in everyday meal. WHO states that malnutrition in all forms includes undernutrition (wasting, stunting, underweight), overweight, obesity, resulting in diet-related non-communicable diseases. Globally, in the year 2020 1.9 billion adults were categorised as obese and 462 million were underweight [14]. A balanced diet helps us to acquire all the required nutrients from food. There are more than 40 varied kinds of nutrients present in food which can be categorized into 7 major groups (carbohydrates, proteins, fats, vitamins, minerals, dietary fibre, water) [15]. In addition to this other essential nutrient categories include (iron, calcium, lipids, sodium, iodine, zinc, magnesium, potassium) [16]. Deficiencies in such key nutrients can weaken the parts of immune system [17].

This paper aims to develop an automated system to recognize and classify food images, and also provide a detailed estimate of nutrient values present in the food item. Traditional methods of dietary assessment include food records, food frequency questionnaires, 24-hour recalls, and screening tools which solely depend on user manual input [18]. Manual methods are subject to errors due to the quality of user input as well as the expertise and knowledge of experts and dietitians in the field. With the advancements in technology and usage of mobile phones with internet facility, a deep learning model which identifies food from images is more reliable and accurate. There are a variety of visual-based dietary assessment techniques available which can be divided into image-assisted and image-based approaches. The former one need more manual intervention where dietitian examines the image physically and presents nutrients value report based on his expertise. The later one is automatic system where no human involvement is needed [19]. However, the implementation of image-based dietary assessment methods is more complicated since it relies heavily on computing algorithms due to its fully automated characteristics. Our work has wide-ranging applications across various domains, including self-diet monitoring for individuals seeking to make healthier choices, support for dietitians in their practice, enhancement of the dining experience at restaurants and eateries, and even provide nutrition education for children.

Our study examines different deep learning models for classifying food images and proposes a best model that requires less training time while delivering a high grade of detection accuracy on the Food-101 dataset. Extensive training and experiments are done on the dataset using various models, ultimately achieving state-of-the-art accuracy of 99.6% with CNNovaNet model architecture. A comparative analysis over three models is presented highlighting the advantages and challenges in implementation. For dietary assessment we used the USDA national food database as a standard source of reference [20].

RELATED WORKS

The existing work shows that researchers have been studying food image recognition realm since last 20 years. Initially machine learning methods were employed in image recognition. Feature descriptors like histogram of gradient, color correlogram, bag of scale-invariant feature transform, local binary pattern, spatial pyramidal pooling, speeded up robust features (SURF), etc, have been applied with some success on small datasets [26]. *Hoashi et al* [27] and *Joutou et al* [28] propose multiple kernel learning methods to combine various feature descriptors. The features extracted have generally been used to train Support Vector Machines (SVM) with a combination of features to boost accuracy.

Deep learning architectures started giving remarkable results since 2014 where images were trained using Convolution Neural Networks [29] achieving highest ever accuracy of 88 %. In the following years scholars explored various deep learning architectures achieving more higher accuracy. *Minki Chun* [4] employed *Inception Resnet* on Korean food image dataset achieving accuracy of 90.2 %. *S. Memis and B. Arsalan* [1] used ensemble network of *ResNeXT and DenseNet* architectures on the UEC FOOD 100 dataset achieving accuracy of 81 %. There is another significant architecture VGG16 introduced by Karen Simonyan [22], which won *ImageNet Large Scale Visual Recognition Challenge (ILSVRC)* in the year 2014. *Landu Jiang and Bojia Qiu* studied a wide used Japanese food dataset the UEC FOOD 256 in 2022 and published various methods for classifying multi-item food images [2]. They employed Region proposal network developed from *faster R-CNN model* to draw bounding boxes across each item in the image and classified the regions using *VGG 16 model* with 90 % accuracy. They also achieved highest accuracy of 93% using VGG model on Food 101 dataset. *Shady Elbassuoni*, classified the UEC FOOD 256 dataset into four NOVA classes stating the state of processing performed on food [3]. Their work focused on healthiness of the food item. *Paritosh Pandey* employed ensemble network of *GoogleNet, AlexNet, and resNet* on the ETH Food dataset and achieved accuracy of 91 % [11].

Deep neural network architectures with many layers achieved much higher accuracy, but this took longer training time. To reduce the time of training a model, techniques like *Transfer Learning* and usage of *Batch Normalization layers* were employed. Transfer learning is a technique where knowledge learned from a previous task is reused in current similar task. *Srikanth Tammina* used transfer learning with *VGG16* architecture on image classification problem and achieved good accuracy of 98 % with few parameters [23]. *Anil and Gangodkar* applied transfer learning technique on three models, *Resnet 50, Xception, and VGG16* for

Pneumonia detection problem [5] where VGG16 model outperformed the other two with 98.27 % accuracy and minimum loss of 4.35%. To reduce the number of parameters in VGG16 model *Xuesong Jin* employed a new model *VGG-S* with 5 convolution layers and 2 fully connected layers [7]. This model worked well for small datasets with significant accuracy.

Employing *Batch Normalization* layers along with VGG16 architecture further reduced the number of parameters and training time for model. *Haizhuang Liu* and *Nontawat Pattanajak* in their works compared the working of VGG16 with and without batch normalization layers and proved that model with batch normalization layers gives stable results, high accuracy, minimum loss and consistent training throughout input data [6][7]. Batch Normalization reduces the internal covariate shift and improves the efficiency of neural networks. By normalizing activations throughout the network, it prevents output distributions and small changes in layer parameters by affecting final output [25].

Nutrition value estimates are done with respect to calorie calculations [10] and presenting top five macro nutrient estimate [2]. A lot of methods were employed to figure out the volume of the food from image such as, using stereo vision cameras with known focal length, using depth sensor cameras, crowdsourcing methods, using additional reference object in the scene while capturing the image, using user thumb as reference image while capturing image, shape template 3D construction method [9]. However, most of the techniques are not suitable to implement outside laboratory settings, where nutrient information varies according to preparation of the food.

FOOD IMAGE DATASETS

There are only a few numbers of food image databases available which vary in number of images and types of items they cover. The cuisines represented in the datasets range across the world. **TABLE 1** lists popular databases of food images which clearly states the number of images, cuisines covered, year of invention, citations done on each dataset [1].

TABLE 1. Food Image Datasets with their classes, cuisines and citations.

Sr No	Dataset	Classes &Images	Cuisine	Year	#Cit
1	PFID	61 / 1098	Fast Foods	2009	254
2	Food-85	85 / 8500	Asian	2010	120
3	UEC FOOD 100	100 / 12740	Japanese	2012	274
4	UEC FOOD 256	256 / 31397	Japanese	2014	170
5	FOOD 101	101 / 101000	Eastern & Western	2014	687
6	UNICT FD889	889 / 3583	Generic	2014	68
7	Vireo FOOD 251	215 / 169673	Chinese	2016	80

The availability of food image datasets has significantly contributed to the field of computer vision, particularly in the domain of food recognition. The first publicly available dataset was Pittsburgh Fast-Food Image Dataset [PFID] was released in the year 2009 which contained food items belonging to 61 different classes. Next widely presented dataset was Food-85 in the year 2010 which included 85 varieties of Asian food items. In 2012 a Japanese database UEC FOOD 100 was introduced which contained 100 food classes captured from different environments. As an extension to this, UEC FOOD 256 database was proposed in the year 2014 which included 31397 images from 256 classes. Both UEC datasets contained multi-food item images where each image has a serving plate with multiple food items.

The same year 2014, witnessed the release of another remarkable dataset Food-101[21], which represented wide range of cuisines from around the world. This dataset comprised of 101 classes of food items with 1000 images per class, which totalled making up 101000 images. Each image is perfectly captured in a clear background featuring only a single food item per plate. As indicated in **TABLE 1**, Food-101 received significantly higher citations compared to others, making it a preferred choice for new experiments. There are two other datasets UNICT FD889 and Vireo FOOD 251 brought out in years 2014 and 2016 respectively. The former one is generic food dataset which had 889 classes of items but a limited number of images per class making up a total of 3583 images. The later one is the latest among food image datasets focusing on Chinese food items from 215 distinct classes.

We conducted our research using Food-101 dataset with the aim of exploring wide cuisines globally. All the images were labelled and divided into particular classes. **Figure 1** shows sample images taken from Food-101 dataset.

Fig 1. Sample images from food 101 dataset

SYSTEM DESIGN

A) DEEP MODELS

The main objective of this work is to develop a deep learning model that requires less training time to classify food images with high accuracy. We started employing CNN architectures for model training and trained the dataset initially with Basic CNN model architecture. CNNs are effective at image classifications because they extract spatial hierarchies of features such as edges, textures and shapes which are important for object recognition. The term ‘convolution’ denotes a mathematical operation where two functions are multiplied to produce a third function which expresses how the shape of one function is modified by the other.

A) BASIC CNN

The basic CNN architecture contains 3 convolutional layers and 2 fully connected layers and max pooling layers. The convolutional layer applies filters to the input image to extract features, the pooling layer down samples the image to reduce computation, and the fully connected layer makes the final prediction. Max pooling layers were added in between convolutional layers to reduce the size of feature maps thereby reducing the complexity in training. This model has *32,69,598* parameters to train.

B) VGG16 MODEL

VGG16 is a standard neural network architecture from Visual Geometry Group developed with 16 layers. It contains 13 convolutional layers and 3 fully connected layers. The convolution layers were divided into 5 blocks where a max pooling layer is added at the end of each block. Max pooling layers are added for down sampling the input. The total parameters for this network would be *2,11,74,238*. After applying transfer learning on VGG16 architecture the total trainable parameters were reduced to *64,59,550*. This decreases the training time for the network. Transfer learning is the improvement of learning in a new task through the transfer of knowledge from a related task that has already been learned [23][24]. **Figure 2** demonstrates VGG16 model architecture with transfer learning applied.

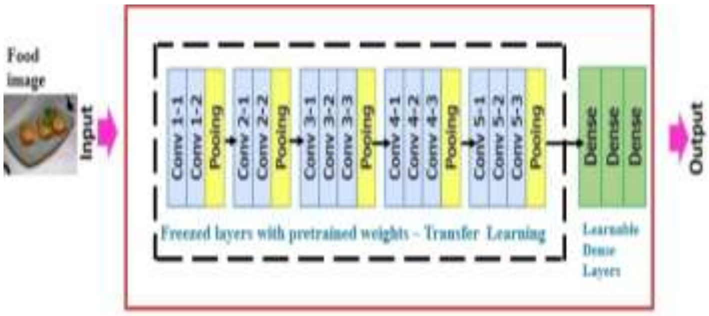

Fig 2. VGG16 model architecture with transfer learning applied

C) CNNovaNET MODEL

To achieve higher accuracy with much less training time we developed a new lightweight CNN model, *CNNovaNet* architecture by applying batch normalization on convolution layers. Batch normalization is a widely used method to reduce the training time on deep learning models. Batch normalization layers prevent large changes that may happen in data distribution, as it reduces internal covariate shift.

BATCH NORMALIZATION

Normalization is a data pre-processing technique which is used to standardize feature values within a dataset to a uniform scale. It serves as the purpose of simplifying data analysis and modelling while mitigating the influence of various scales on deep learning models. *Batch Normalization* is a deep learning approach particularly used to improve efficiency and reliability of neural networks [8]. It is a supervised learning method for normalizing interlayer outputs of a neural network. As a result, the next layer receives a 'reset' of the output distribution from the preceding layer, allowing it to analyse the data more effectively. Input standardizations will decrease the dropout rate, which leads to vast increase in precision.

Internal Covariate Shift is the change in the distribution of network activations due to the change in network parameters during training [25]. With batch normalization, *Internal Covariate Shift* can be reduced effectively. **Figure 3,** illustrates the working of batch normalization layer.

Input: Values of x over a mini-batch: $\mathcal{B} = \{x_{1...m}\}$
Parameters to be learned: γ, β
Output: $\{y_i = BN_{\gamma,\beta}(x_i)\}$

Mini-batch mean: $$\mu_\beta \leftarrow \frac{1}{m}\sum_{i=1}^{m} x_i$$

Mini-batch variance: $$\sigma_\beta^2 \leftarrow \frac{1}{m}\sum_{i=1}^{m}(x_i - \mu_\beta)^2$$

Normalize $$\hat{x}_i \leftarrow \frac{x_i - \mu_\beta}{\sqrt{\sigma_\beta^2 + \epsilon}}$$

Scale and shift: $$y_i \leftarrow \gamma \hat{x}_i + \beta \equiv BN_{\gamma,\beta}(x_i)$$

Fig 3. Algorithm of Batch Normalization, applied to activation x over a minibatch

CNNovaNet is a custom lightweight model inspired from the architectures of basic CNN and VGG16. The model has 7 layers in the network, 5 convolutional layers and 2 fully connected dense layers. In addition, 4 max pooling layers and 3 batch normalization layers were added in between convolution layers for better performance of the network. **Figure 4** illustrates the architecture of CNNovaNet model.

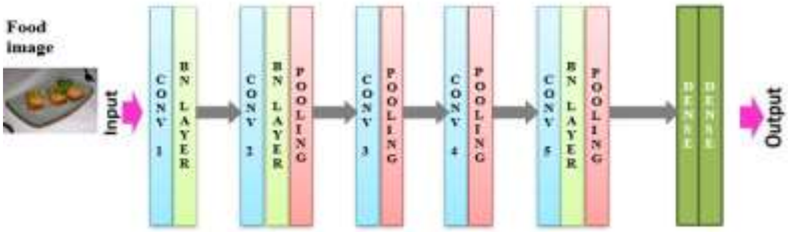

Fig 4. CNNovaNet model architecture

The model accepts input images with size (128,128,3), maintaining the kernel size as 3 x 3 across the network. First convolution layer (*conv 1*) has 32 neurons, it is followed with a batch normalization (*BN layer*). The activations from BN are given to *conv 2* layer with 32 neurons. *Conv 2* is followed by *BN layer* and *max pooling* layer. Both *conv 3* and *conv 4* has 64 layers and are immediately followed by a pooling layer. Finally, *conv 5* is the last convolution layer with 64 neurons, accompanied by 1 *BN layer* and 1 pooling layer. The last activation outputs are flattened and forwarded to dense layers. The activation function employed is *'sigmoid'* activation, and optimizer is configured as *'rmsProp'*. Total parameters of the network are 6,31,038.

EXPERIMENT AND RESULTS

We implemented our deep learning models using Keras, which is a high-level neural network API written in python. In the backend we used TensorFlow platform designed by Google. The entire training was done on a GPU [Graphical processing unit] machine, NVIDIA TITAN XP GRAPHICARD with 3840 CUDA cores. It is equipped with a powerful Intel Core i7-8700K processor, and paired with an ASUS Z370 motherboard to ensure efficient communication between components. It is a high-performance graphics processing unit designed for computationally intensive tasks which gives excellent support for high-quality rendering and AI-driven applications.

We selected 30 most common foods from the dataset Food-101 and made our training data ready with 30 classes and 200 images per class, totalling 6000 images. This dataset was trained over 3 different models, where each model is evaluated against accuracy and training time. A comparison analysis is made stating the advantages and challenges in implementing the models.

A) Basic CNN Model

The model is designed to accept images with dimensions of (128,128,3). All the input images were resized to required size as part of data pre-processing. Training was done using '*Relu*' activation function and '*adam*' optimizer. The model was trained over 20 epochs where it converged around 17 epochs with loss 0.3689 and 89% accuracy. **Figure 5 (a)** shows the accuracy curve and, **Figure 5 (b)** shows loss curve for basic CNN model training.

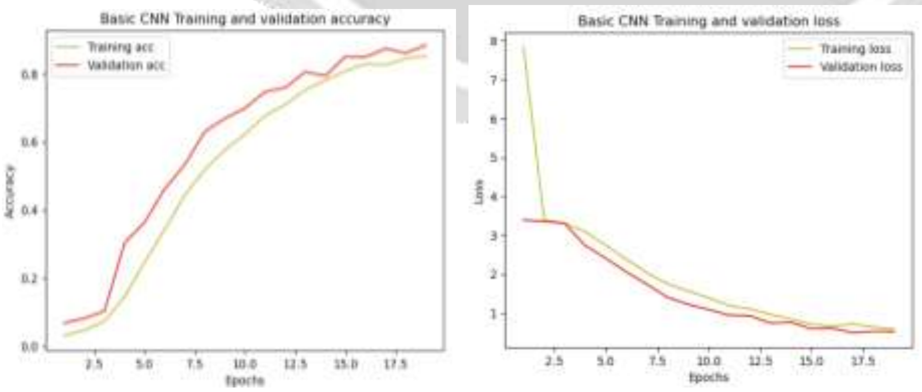

Fig 5. A) Training and validation ,accuracy curve of basic CNN model(left)

B) Training and validation ,loss curve of basic CNN model(right)

B) VGG 16 Model

The model is designed to accept images with dimensions of (224,224,3). All the input images were resized to required size as part of data pre-processing. Training was done using '*Relu*' activation function and '*adam*' optimizer. The model was trained over 20 epochs where it converged around 8 epochs with loss 0.0884 and 99.67 % accuracy. **Figure 6 (a)** shows the accuracy curve and, **Figure 6 (b)** shows loss curve for VGG 16 model training

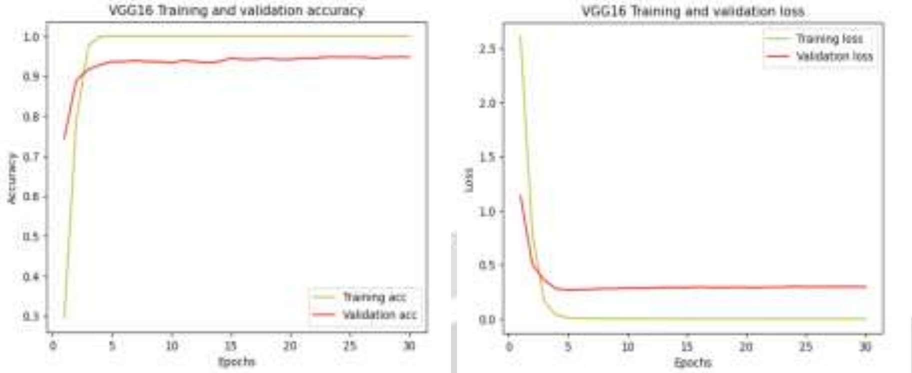

Fig 6. A) Training and validation ,accuracy curve of VGG16 model(left)

B) Training and validation ,loss curve of basic VGG16 model(right)

C) CNNovaNET MODEL

The model is designed to accept images with dimensions of (128,128,3). All the input images were resized to required size as part of data pre-processing. Training was done using '*sigmoid*' activation function and '*rmsProp*' optimizer. The model was trained over 20 epochs where it converged around 19 epochs with loss 0.1829 and 98.33 % accuracy. **Figure 7 (a)** shows the accuracy curve and, **Figure 7 (b)** shows loss curve for basic CNNovaNet model training

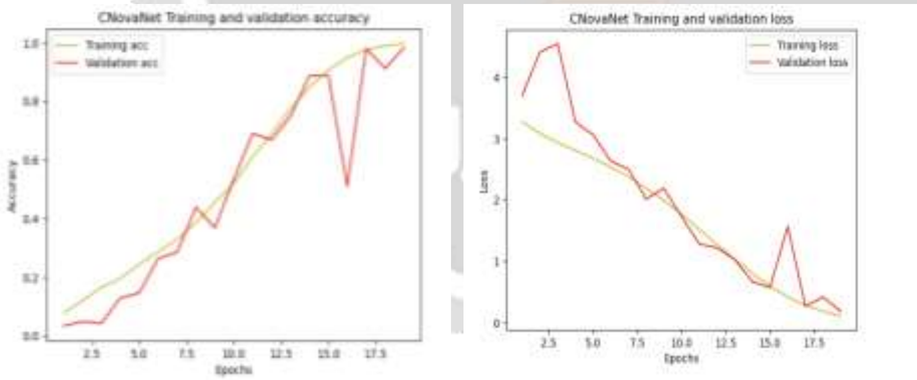

Fig 7. A) Training and validation ,accuracy curve of CNNovaNet model (left)

B) Training and validation ,loss curve of CNNovaNet model (right)

COMPARISION OF MODELS

On comparing three deep learning models, it's evident that VGG16 outperformed the other two models with an impressive accuracy of 99.67%. VGG16, being a standard deep learning model proved for image classification, performed extremely well on our dataset with high accuracy. But, due to the complexity in model architecture, the number of parameters in VGG 16 were too high even after applying transfer learning. As we can see from the table below, VGG16 had 64,59,550 parameters which took more time for execution. In contrast, the light weight model CNNovaNet gave exceptional results employing much fewer parameters 6,31,038. On comparing the training time of both networks, we could clearly state CNNovaNet takes less time for training, achieving accuracy of 98.33 % which was very close to that of VGG16. The basic CNN model, while faster in execution yielded lower accuracy because of its architectural design. Finally, to meet our objective of training a model in less time giving high accuracy CNNovaNet emerges as a preferrable choice. **Table 2** provides detailed analysis of the three models.

TABLE 2. Comparison of implementations and results of 3 deep learning models (Basic CNN, VGG16, CNNovaNet)

Sr No	Category	Basic CNN	VGG 16	CNNovaNet
1	Layers	5 3 Convolution layers 2 Fully connected layers	16 13 convolutional layers 3 Fully Connected Layers	7 5 Convolutional Layers 2 Fully connected layers
2	Number of classes	30	30	30
3	Number of images per class	200	200	200
4	Size of each image	(128,128,3)	(224,224,3)	(128,128,3)
5	Optimizer	Adam	Adam	RMSProp
6	Activation Function	ReLu	ReLu	Sigmoid
7	Epochs	20	20	20
8	Transfer Learning	NO	YES	NO
9	Number of parameters trained	32,69,598	64,59,550	6,31,038
10	Training time	50 seconds	180 seconds	83 seconds
11	Accuracy	89.1 %	99.67 %	98.33 %
12	Loss	0.3689	0.0884	0.1829

DIETARY ASSESSMENT

The Agricultural Research Service, a branch of the U.S. Department of Agriculture (USDA), is dedicated to advanced research in food science and nutrition. One of their significant contributions is the maintenance and continuous updating of a comprehensive database that encompasses a wide range of common food varieties found globally [20]. This extensive dataset provides detailed information on over 100 different nutrients, including both macro and micro nutrients, present in each food item. To facilitate accessibility each food item is represented in a separate Json format file.

We have extracted the top ten most essential nutrient types like (protein, calcium, carbohydrates, fats, vitamins, iron, fibre, calories, water, sugars) for all food classes in Food-101 dataset and used for our dietary assessment. Each food item in the dataset includes a '*servingSize*' parameter, which serves as the basis for quantifying nutrient information. **TABLE 3** illustrates a sample of food nutrient values extracted from USDA database.

TABLE 3. Nutrient values for sample food items taken from USDA database

Class Label	Name	Protein	Calcium	Fat	Carbohydrates	Vitamins	Energy	Water	Sugars	Fiber	Iron	ServingSize	ServingSizeUnit
9	breakfast burrito	8.13	7	7.77	26.2	534.2	198	81.7	1.41	1.4	1.27	283	g
10	bruschetta	4.29	50	7.61	21.3	537.9	171	65.6	3.34	1.5	1.58	32	g
11	caesar salad	4	100	16	8	4506	190	65.6	2	1	0.72	100	g
12	cannoli	6.32	21	14.7	42.1	0	328	65.6	39	0	0.95	95	g
13	caprese salad	5.76	79	10.5	9.42	0	152	65.6	4.71	1	0.94	191	g
14	carrot cake	3.7	0	33.3	59.3	0	519	65.6	55.6	0	0	27	g
15	ceviche	10.5	22	0.76	3.58	18.6	63	84	1.34	0.5	0.16	100	g
16	cheese plate	6.46	146	11.5	15.6	172.6	192	64.8	2.83	0.7	0.55	28	g
17	cheesecake	7.02	35	19.3	36.8	0	351	64.8	21	1.8	0.63	57	g
18	chicken curry	6.47	20	6.45	6.47	8.6	106	79.1	1.6	1.4	0.73	140	g
19	chicken quesadilla	10.6	133	5.31	15	266.1	150	79.1	0.88	5.3	0.96	113	g

To present this data to users according to their serving size, we use the *NutriScaler* algorithm, ensuring accurate and user-specific dietary assessments. In real time environment any user can capture their food image using their smartphone and log it into our application specifying the size of the food as additional parameter. Our deep learning model predicts the type of food from the image and gives the output label (X_i) to *NutriScaler algorithm.* The algorithm receives two inputs, one is food label (X_i) and the other is serving size value (W) from the application. The scaling function $X_i\,[Y_{m'}]$ is applied over all the 10 nutrient categories (Y_0 to Y_9) belonging to X_i and new values ($Y_{m'}$) are computed and presented to user as a report. The system gave accurate results for all food items. **Figure 8** shows mathematical function of NutriScaler algorithm. The function can be applied for all food items in Food-101 dataset.

$$X_i\,[Y_{m'}] = X_i\,[Y_m] * W \;/\; Xi[S]$$

X_i	= Food Categories, where i ranges from 0 to n (n food items)
Y_m	= Nutrient value, where m ranges from 0 to 9 (10 different nutrients)
S	= Standard serving size parameter
$X_i[S]$	= Serving size of i^{th} food item
W	= New serving size value provided by user
$Y_{m'}$	= Scaled nutrient values with respect to W

Fig 8. NutriScaler function

CONCLUSION AND FUTURE SCOPE

Our study focuses on the classification of processed food items based on images and the determination of various nutrient values present in that food. We explored the performance of three deep learning models, namely the basic CNN, VGG16, and CNNovaNet on the Food101 dataset. With a view to fasten the training time of our model and yielding highest detection accuracy, much care was taken on the number of parameters we train. CNNovaNet model delivered outstanding results across the dataset by achieving high accuracy in limited time, making it the best model for our dataset. VGG16 exhibited similar accuracy but consuming more time for training. Basic CNN model however took considerable time for execution but gave less accuracy compared to the other two models. We successfully extracted ten nutrient types for each food item and presented to user along with scaled values according to user provided serving size. Looking ahead, there are ample opportunities for expanding this work to encompass datasets featuring diverse cuisines, particularly Indian cuisine, which is currently underrepresented. Furthermore, a potential area for improvement includes quantifying the amount of food from user-provided images and making precise predictions regarding nutrient content based on actual consumption.

REFERENCES

[1] Arslan,B., Memis,S., Sonmez,E.B., Batur,O.Z.(2022). Fine-Grained Food Classification Methods on the UEC FOOD-100 Database. Volume 3. IEEE Transactions on Artificial Intelligence.

[2] Jiang,L., Qiu,B., Liu,X., Huang,C. (2020). DeepFood : Food Image Analysis and Dietary Assessment via Deep Model. Volume 8. IEEE Access.

[3] Elbassuoni,S., Ghattas,H., El Ati,J. (2022). DeepNOVA: A Deep Learning NOVA Classifier for Food Images. Volume 10. IEEE ACCESS.

[4] Chun,M., Jeong,H., Lee,H., Yoo,T., Jung,H. (2022). Development of Korean Food Image Classification Model Using Public Food Image Dataset and Deep Learning Methods. Volume 10.IEEE ACCESS.

[5] Anil,B.D., Gangodkar,D.R., (2022). Comparative Analysis Using Transfer Learning Models VGG16, Resnet50, Xception to Predict Pneumonia. 2nd International Conference on Innovative Sustainable Computational Technologies.

[6] Liu,H., Gong,H., Ding,X. (2021). Food Image Recognition Algorithm Base on Improved VGG16. IEEE 2nd International Conference on Information Technology,Big Data and Artificial Intelligence.

[7] Jin,X., Du,X., Sun,H. (2021). VGG – S: Improved Small Sample Image Recognition Model based on VGG 16. 3rd International Conference on Artificial Intelligence and Advanced Manufacture.

[8] Pattanajak,N., Malekmohamadi,H. (2019). Improving a 3-D Convolutional Neural Network Model Reinvented from VGG16 with Batch Normalization. IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing &Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation.

[9] Ahmed Subhi,M., Hamid Ali,S., Abulameer,M. (2019). Vision-Based Approaches for Automatic Food Recognition and Dietary Assessment: A Survey. Volume 7. IEEE ACCESS.

[10] Turmchokkasam,S., Chamnongthai,K. (2018). The Design and Implementation of an Ingredient-Based Food Calorie Estimation System Using Nutrition Knowledge and Fusion of Brightness and Heat Information. Vokume 6. IEEE ACCESS.

[11] Pandey,P., Deepthi,A., Mandal,B., Puhan,N.B. (2017,December). FoodNet: Recognizing Foods Using Ensemble of Deep Networks. Volume 24. IEEE SIGNAL PROCESSING LETTERS.

[12] Thompson.F.E., Byers,T. (1994). Dietary Assessment resource manual. National Library of Medicine. United States. https://pubmed.ncbi.nlm.nih.gov/7965210/

[13] Wow Essays. (2020,March,3). Obesity Report Samples. https://www.wowessays.com/freesamples/obesity-report-samples/.

[14] World Health Organization. (2021, June 9). Malnutrition. https://www.who.int/health-topics/malnutrition#tab=tab_1.

[15] Hong Kong Centre for Health Protection.(2017, December 4). Health Topics.Nutrient Classifications. Hong Kong Government https://www.chp.gov.hk/en/static/100022.html.

[16] National Center for Biotechnology Information. (2023, May 1). National Library of Medicine. United States Government https://www.ncbi.nlm.nih.gov/books/NBK554545/.

[17] National Health Information Service (NHS inform). (2023, January 4). Health benefits of eating well. Scotland Government https://www.nhsinform.scot/

[18] Bailey, R. L. (2021,March). Overview of dietary assessment methods for measuring intakes of foods, beverages, and dietary supplements in research studies. Current opinion in biotechnology.

[19] Lo.F., Qiu,J. (2020). Image-Based Food Classification and Volume Estimation for Dietary Assessment: A Review. Volume 24. IEEE Journal of Biomedical and health informatics.

[20] US Department of Agriculture, Agricultural Research Service. (2018). Nutrient Data Laboratory. http://www.ars.usda.gov/nutrientdata

[21] Bossard,L., Guillaumin,M., Van Gool,L. (2014). Food-101–mining discriminative components with random forests. Proc. European conference on computer vision.

[22] Simonyan,K., Zisserman,A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv preprint arXiv:1409.1556*.

[23] Srikanth.T, (2019,October). Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images. Volume 9. International Journal of Scientific and Research Publications.

[24] Kaur,T., Tapan Kumar,G., (2019,December). Automated Brain Image Classification Based on VGG-16 and Transfer Learning. International Conference on Information Technology.

[25] Ioffe,S., Szegedy,C. (2015). Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. Proceedings of Machine Learning Research. https://proceedings.mlr.press/v37/ioffe15.html.

[26] Martinel,N., Piciarelli,C., Micheloni,C. (2016). A supervised extreme learning committee for food recognition. Volume 148. Comput, Vis.Image Understanding.

[27] Hoashi,H., Joutou,T., Yanai,K. (2010,December). Image recognition of 85 food categories by feature fusion. Proc. IEEE Int. Symp Multimedia.

[28] Joutou,T., Yanai,K. (2009,November). A food recognition system with multiple kernel learning. Proc. IEEE Int. Conf. Image Process.

[29] Kawano,Y., Yanai,K. (2014). Food image recognition with deep convolutional features. International Joint Conference on Pervasive and Ubiquitous Computing.