Grenze International Journal of Engineering and Technology, January Issue

Food Detection & Calorie Estimation of Food & Beverages using Deep Learning

Anurag Mishra[1], Ankur Mishra[2] and Ayush Gupta[3]
[1-3]ABES Engineering College, Ghaziabad, India
Email: anuragmishra7329@gmail.com, ankurmisra01@gmail.com, ayush.gkp.100@gmail.com

***Abstract*—People nowadays, are not maintaining a healthy body due to intake of unnecessary and unhealthy diet. Hence, it is required in recent times, to maintain their standard diet so they can avoid various problems related to obesity that cause risk to their healthy body. So in this paper, we are presenting an automatic machine learning-based novel system that classifies the food items, identifies them, and estimates the calories present in them i.e. estimation of calories is performed. We have obtained an accuracy of 89.48%. For the classification of the food images and identification, we suggest using a Convolutional Neural Network based on a deep learning model that will perform the required purpose in the training part of the system. The proposed model would utilize a client- server architecture, wherein the client uploads an image. A needed algorithm on the server side then executes to assist in estimating the caloric content of the corresponding food items. Here we are using the dataset of thousands of images from Kaggle and applying the appropriate algorithm to predict the estimation with the best accuracy. Here we will be using CNN and then pre- trained CNN models such as MobileNet, ResNet, etc.**

***Index Terms*— Calorie Estimation, MobileNet, EfiicientNet1.**

I. INTRODUCTION

Nowadays, people are showing keen interest in calorie measurement because people are measuring weight, maintaining a healthy diet, and staying away from obesity to keep them fit and less prone to diseases. Adult obesity rates are rising at a startling rate. The primary factor contributing to obesity is the imbalance between energy from food and the amount of nutrition consumed. A diet heavy in calories can be detrimental and aggravate many different conditions. Breast, colon, and prostate cancer are brought on by a high calorie diet. The second major cause of cancer is a high-calorie diet. According to dietetics, the human body needs a specific amount of calories in order to maintain the right calorie balance. Over ten percent of adults worldwide are obese, according to the World Health Organization. The medical condition known as obesity is characterized by an accumulation of extra body fat that may be dangerous to one's health. A person is considered obese if their daily caloric intake is greater than their daily energy expenditure. Being overweight is associated with obesity. The American Medical Association formally classified obesity as a condition in 2013 with significant health consequences, necessitating medical intervention [2].

Obese persons cannot lose weight more healthfully or healthy people grow healthier unless daily food consumption is consistently measured. The study of the user's food intake log for the preceding 24 hours is the basis of the traditional technique. Although the clinical display has certain benefits, but these techniques usually

Grenze ID: 01.GIJET.11.1.420

make the user forget about the patient's agony or make them announce that they don't want to utilize the programs. These days, smart devices are becoming more and more necessary in daily life, and their usage is growing. Devices for the treatment of various disorders are not uncommon. In order to accomplish this, we propose a procedure or tool that assists both normal and obese individuals in maintaining a balanced diet by tracking the types and amounts of food they eat each day whenever it's convenient for them. By providing a picture of the meal system, the proposed idea will enable the user to ascertain the food item's composition. Using a Convolutional Neural Network crawl, the application will identify and detect the food items in the picture, Additionally, the system will be able to estimate food quality information from the web. The suggested method will enable healthy individuals as well as those who are fat to plan their daily caloric intake ahead of time. The ways we will expand on this thesis are as follows:.

- We present a novel approach based on transfer learning that computes food attributes and precisely classifies photos of food.
- In order to assess the present system and upcoming deep learning-based recognition systems, we present the dataset

II. Literature Review

Numerous studies on the process of identifying food have been conducted using a variety of methodologies. But they are all limited in different ways. We have thoroughly examined a few of the previously suggested models for food identification and calorie estimation. In the research paper by Shahriar Ahmed Ayon, Chowdhury Zerif Mashraf, Abir Bin Yousuf, Fahad Hossain, Muhammad Iqbal Hossain, they have created a brand-new neural network-based model that can identify food items in an image and provide us with an estimate of how many calories each item contains. They have put together a collection of about 23000 photos for 23 distinct cuisine categories in order to accomplish the goal. In order to do this, they have developed the system that uses CNN's properties that were taken from Inception V3 to train it to recognize various cuisines. The technique was implemented on a webpage where users input images of food items to have the system guess the food item and its projected calories in real time. They have achieved 89.48% accuracy with this model.

In Haoyu Hu, Zihao Zhang and Yulin Song 2020 research paper, a model for estimating the calories in the Chinese food have been developed. They employed object detection to assess the calorie count of some well-known Chinese foods as well as Western dishes in an effort to voice their worries about this matter and demonstrate their keen interest. They created several image-based calorie estimation models based on the food photos and previously defined calorie data in the hopes that these models would correctly identify the names of Chinese and Western foods, provide information on their calorie content and recipes, and, in the end, suggest meal plans for various population groups. They employed the SSD (Single Shot Multi Box Detector), which processes item detection and classification in real time, to identify the dish.

In Takumi Ege and Keiji Yana research paper, they put food photographs of several cuisines using Faster R-CNN, a popular object detection technique. They include two different types of food photo datasets in the studies to verify. Additionally, food photographs of various dishes are used to estimate calories using this food detector.

III. Recommended Method

A. Compiling the Dataset

The FOOD101 dataset, which is available on Kaggle, was utilized for training purposes, here there are 101 food categories 1,01,000 photos in this dataset. There are 750 training photos and 250 manually approved test images available for each class.

The Nutrition 5k dataset on Kaggle contains detailed nutritional information for around 5,000 food items, covering various categories like fruits, vegetables, grains, meats, dairy products, and processed foods. Each item includes data on calories, proteins, fats (including types), carbohydrates (including fiber and sugars), vitamins, minerals, cholesterol, and other nutrients. The dataset is provided in CSV format, making it accessible for analysis and use in diet planning, nutritional studies, and machine learning projects related to food and health. It serves as a comprehensive resource for anyone interested in nutrition science or developing health-related applications.

The Recipes 5k dataset on Kaggle includes around 5,000 recipes, each with a list of ingredients, instructions, and nutritional information. The dataset covers various cuisines and meal types, providing details such as ingredient quantities, cooking steps, and estimated nutritional values like calories, proteins, fats, and carbohydrates. Provided in CSV format, the dataset is ideal for use in culinary research, meal planning, recipe recommendation systems, and machine learning projects focused on food and health. It offers a rich resource for exploring and analyzing culinary data.

B. Sample dataset

Both of our datasets were created using a total of 101 distinct food categories, each of which has 1000 photos. It is therefore exceedingly challenging to display a high-quality sample from the database. We attempted to display few photographs from our dataset, and we were able to compile in Figure1. This sample image below shows that we employed a variety of food types and calorie counts when creating our dataset. It also includes few photographs from each meal category.

Fig.1 A Few Sample Pictures from the Collection

C. Dataset Preprocessing

The convolutional neural network, or CNN primarily uses pooling to minimize the dimension, is needed to analyze the data. CNN now attenuates the image into column vectors after preparing the images for multilayer perceptrons. Using back propagation, we may backtrack to analyze and find mistakes once these feed-forward images are placed into a neural network. This process is done for each iteration in training. Once the mistake has been calculated, the back propagation traverses back to the inner or hidden layer to alter the weight to reduce the error. This step is repeated by back propagation till the true correct observation is obtained.

D. Feature Extraction Technique

MobileNetv3 is a family of lightweight, efficient deep neural network architectures designed by Google, particularly optimized for mobile and embedded vision applications. Key features include:

- Efficiency: Uses depthwise separable convolutions to reduce the number of parameters and computational cost compared to traditional convolutions.
- Scalability: Provides a trade-off between latency, accuracy, and model size through adjustable hyperparameters like width multiplier and resolution multiplier.
- Performance: Achieves a good balance between accuracy and speed, making it suitable for real-time applications on devices with limited computational power.

MobileNet is widely used in tasks such as object detection, image classification, and face recognition, especially in mobile and edge computing environments.

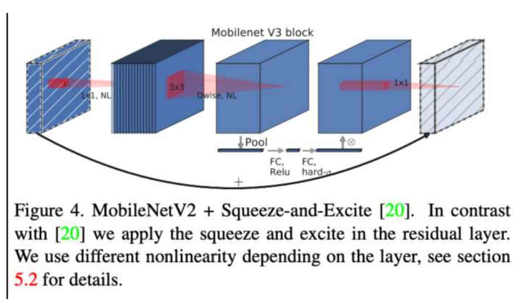

Fig.2 MobileNetV3 Architechture

NASNet Mobile is a deep learning architecture designed by Google for mobile and embedded vision applications. Key features include:

- AutoML Design: Created using Neural Architecture Search (NAS), an automated process to find the most efficient and high-performing architecture.
- Efficiency: Optimized for mobile devices, balancing accuracy and computational efficiency to ensure fast inference times with limited resources.
- Scalability: Offers different model sizes, allowing developers to choose the right balance between performance and computational cost for their specific use case.

NASNet Mobile is particularly suited for tasks such as image classification and object detection on smartphones and other devices with constrained computational capabilities.

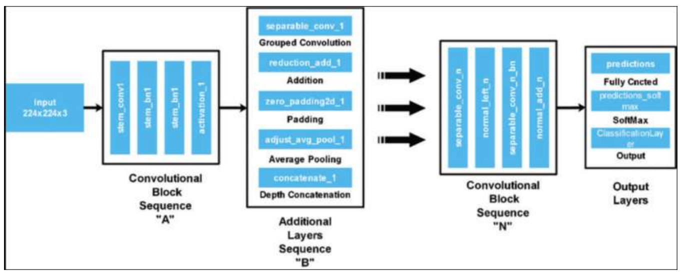

Fig.3 NASNet Mobile Architechture

EfficientNetB1 is a deep learning model that is part of the EfficientNet family, developed by Google to achieve high performance with optimized computational efficiency. Key features include:

- Scaling: Utilizes a compound scaling method to balance network depth, width, and resolution, enhancing both accuracy and efficiency.
- Performance: Provides a good trade-off between accuracy and resource usage, suitable for various vision tasks.

Applications: Ideal for image classification and other computer vision tasks in environments with limited computational resources.

EfficientNetB1 is designed to offer strong performance while minimizing computational and memory demands, making it suitable for deployment on devices with constrained resources.

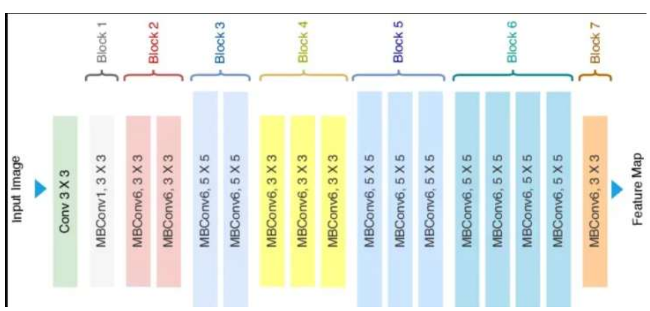

Fig.4. EfficientNet B1 Architechture

IV. MODEL SUMMARY

A work-flow diagram can be used to explain our entire work process. Figure 5's work flow diagram allows us to give a summary of our efforts and the procedures we took to construct our system. These are our code's primary processes; lesser procedures are included in these primary procedures even though they aren't displayed here.

A. Constructing the dataset

Gather and organize the data, ensuring it includes diverse and representative samples for the task. Split the dataset into training, validation, and test sets.

B. Data Preprocessing

Clean and normalize the data, resize images, and apply augmentations to improve model robustness.

C. Building Different Model

Choose and configure MobileNet, NASNet Mobile, and EfficientNetB1 architectures, adjusting hyperparameters for optimal performance.

D. Extracting Features

Use the pre-trained models to extract features from the input data, leveraging transfer learning for improved accuracy.

E. CNN Training

Fine-tune the models on the training dataset, using the extracted features, with appropriate optimization techniques to enhance performance

F. Save Model

Save the trained models to disk, including architecture, weights, and configuration for future inference and deployment.

G. Frontend making

Develop a user-friendly frontend interface to allow users to interact with the models, input data, and view predictions, integrating the saved models for real-time usage.

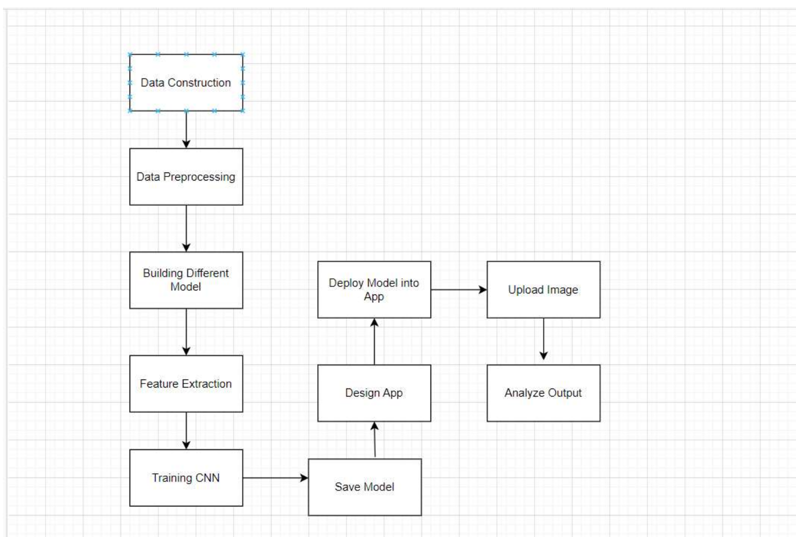

Fig.5. Diagram of Proposed System Workflow

V. CONCLUSION AND FUTURE WORK

The primary goal of this research is to determine a food's calories from an image, which will aid in the treatment of overweight-related conditions such as diabetes, obesity, heart disease, and renal failure. We think that by keeping track of their calorie intake, individuals who are aware of food's calories will be able to lead better lives. Thus, we have investigated and studied a number of approaches for the process of identifying food. CNN works best for tasks involving object or picture detection.They also offer high calculation efficiency for object or picture classification systems when it comes to statistical outcomes.

At this point in the study, we concentrated on developing a CNN algorithm-based webpage utilizing ResNet and Inception v3 models, which can categorize various food kinds and display their nutritive worth. We found that our system has the following shortcomings: Food items with similar appearances pose a challenge for multiple food detectors, as our system may identify numerous food items in an image with just one food item. Finally, the angle at which the picture is captured is critical to our system's ability to correctly identify food items. Our system typically performs badly for photos that have top view.The frontend-app is shown in the figure 6.

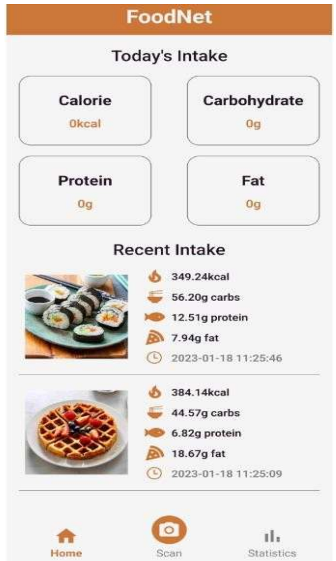

Fig.6. Full App View

REFERENCES

[1] Haoyu Hu, Zihao Zhang, Yulin Song, "Image Based Food Calories Estimation Using Various Models of Machine Learning," in 2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE), IEEE, 2020.

[2] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D.Erhan, V.Vanhoucke, and A.Rabinovich, "Going deeper with convolutions," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 1–9.

[3] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEEconferenceoncomputervisionandpatternrecog- nition, 2016, pp. 770–778.

[4] C. Szegedy, V.Vanhoucke, S.Ioffe, J.Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," inProceedings of the IEEE confer- enceoncomputervisionandpatternrecognition,2016, pp.2818–2826. [5]P.G. Clem, M. Rodriguez, J.A. Voigt and C.S. Ashley, U.S. Patent 6,231,666. (2001). https://doi.org

[5] E. J. Gallagher and D. LeRoith, "Obesity and diabetes: The increased risk of cancer and cancer-related mortal- ity,"Physiologicalreviews,vol.95,no.3,pp.727–748, 2015.

[6] M. A. Subhi and S. M. Ali, "A deep convolutional neural network for food detection and recognition," in 2018IEEE-EMBSConferenceonBiomedicalEngineer- ing and Sciences (IECBES), IEEE, 2018, pp. 284–287.

[7] R. Balsys, Convolutional neural networks (cnn) ex- plained step by step, Feb. 2020. [Online]. Available: https://medium.com/analytics-vidhya/convolutional- neural - networks - cnn - explained - step - by - step - 69137a54e5e7.

[8] S.Saha,Acomprehensiveguidetoconvolutionalneural networks-the eli5 way, Dec. 2018. [Online]. Available: https://towardsdatascience.com/a-comprehensive- guide-to-convolutional-neural-networks-the-eli5-way- 3bd2b1164a53.

[9] Shahriar Ahmed Ayon,Chowdhury Zerif Mashraf,Abir BinYousuf,Fahad Hossain,Muhammad Iqbal Hossain, "FoodieCal: A Convolutional Neural Network Based Food Detection and Calorie Estimation System,"in 2021 National Computing Colleges Conference (NCCC).