ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 11 Issue 08 August 2024

Automated Food Recognition and Calorie Estimation using Deep Learning

[1] Banusharath K A, [2] Karan Karthick R, [3] Karthiban B, [4] Manimegalai C T

[1][2][3] B.Tech Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai, India

[4] Professor, B.Tech Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai, India

Email: [1] bk6524@srmist.edu.in, [2] kr6863@srmist.edu.in, [3] kb8433@srmist.edu.in, [4] manimegc@srmist.edu.in

Abstract— ***In the realm of dietary analysis and nutritional monitoring, the utilization of Convolutional Neural Networks (CNNs) has emerged as a pivotal advancement. This paper presents a thorough investigation into the advancements and breakthroughs achieved through the application of CNNs in the realms of food recognition and calorie estimation. The burgeoning significance of automated dietary analysis necessitates a robust framework capable of accurately identifying food items from images and estimating their calorie content, thereby facilitating informed dietary choices and promoting overall well-being. CNNs exhibit remarkable proficiency in discerning intricate patterns within food images, enabling real-time examination and monitoring of dietary consumption. Furthermore, the integration of calorie estimation capabilities within CNN architectures empowers individuals with vital nutritional information, fostering a deeper understanding of their dietary patterns and facilitating adherence to balanced nutrition guidelines. Despite inherent challenges, including variations in food presentation and image quality, the utilization of CNN algorithms showcases immense potential in revolutionizing dietary analysis and empowering individuals to make informed decisions regarding their health and wellness. This research not only contributes to the advancement of computer vision techniques but also aligns with Sustainable Development Goal 3 – Good Health and Well-being, by promoting informed dietary choices and supporting the attainment of balanced nutrition and calorie management objectives.***

Index Terms— ***CNN, MiDaS, OpenCV, Calorie Estimation, Food Recognition***

I. INTRODUCTION

In recent years, the convergence of deep learning methodologies and dietary analysis has emerged as a focal point within the realms of artificial intelligence (AI) and healthcare[1]. The contemporary lifestyle often characterized by suboptimal dietary habits and limited nutritional awareness has underscored the imperative for innovative solutions facilitating informed dietary decisions and fostering overall well-being. This imperative has spurred the development of sophisticated systems adept at automating food recognition and calorie estimation, thereby empowering individuals to monitor their nutritional intake with heightened precision and efficacy[2].

Deep learning, a prominent subfield of machine learning, has assumed a central role in this endeavor, leveraging its inherent capacity to autonomously discern complex patterns from vast datasets[3]. Through the utilization of deep neural networks, particularly Convolutional Neural Networks (CNNs), researchers and practitioners have endeavored to revolutionize the landscape of dietary analysis by enabling real-time food recognition and calorie estimation directly from visual inputs[4].

The essence of deep learning lies in its multi-layered neural architectures inspired by the structural and functional intricacies of the human brain. This paradigmatic framework has exhibited remarkable proficiency in extracting nuanced features from diverse data modalities, spanning images, audio, and text. Such capabilities have precipitated transformative breakthroughs across a spectrum of domains, encompassing computer vision, natural language processing, and speech recognition[5].

This paper endeavors to furnish a comprehensive review of the latest methodologies and advancements in food recognition and calorie estimation facilitated by CNN algorithms. Drawing upon insights and perspectives from leading experts and professionals in the field, this research seeks to elucidate the fundamental tenets of deep learning in the context of dietary analysis and nutritional monitoring. Furthermore, it endeavors to elucidate the practical applications and broader implications of automated food recognition systems in fostering healthier dietary habits, facilitating weight management, and augmenting overall nutritional literacy.

Subsequent sections will delve into the foundational principles of deep learning, expound upon the manifold applications of CNNs in food recognition, and dissect the advantages and challenges attendant to the integration of deep learning methodologies in dietary analysis. Additionally, the delineation of objectives and motivations underlying this research endeavor will lay the groundwork for a holistic exploration of the transformative potential inherent in deep learning for engendering proactive approaches to health and well-being through informed dietary decision-making.

IFERP Explore Your Research Journey.
ISSN (Online) 2394-2320
International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Volume 11 Issue 08 August 2024

In summary, the intersection of deep learning and dietary analysis represents a paradigmatic shift towards leveraging cutting-edge AI techniques to address contemporary challenges in nutrition and wellness. Through meticulous examination and synthesis of pertinent research and insights, this paper aims to contribute to the scholarly discourse surrounding the utilization of deep learning in fostering healthier lifestyles and promoting optimal nutritional outcomes.

II. RELATED WORKS

In the realm of food recognition and calorie estimation, deep learning stands out as a superior approach, offering enhanced outcomes. This literature review examines the use of deep learning models in the recognition of food items and the estimation of their calorie content. The research involves developing algorithms to identify food and estimate its volume in order to determine its calorie content. Sombutkaew et al. [6] proposed the implementation of a calorie estimation mechanism within an android mobile application. Calorie estimation is conducted by analyzing a food image taken with a smartphone camera and the depth image obtained from the AR core library. The food area is segmented using a refined Mask R-CNN with a dataset of Thai cuisine images. Machine learning techniques such as Linear Regression, Support Vector Regression, K-Nearest Neighbor, and Deep Neural Network are employed to predict the amount of calories in a meal from each photograph. From this, we have learned that Deep Neural Network provides superior prediction outcomes with the highest level of accuracy, the lowest rate of error, and the highest R-Square score. Rahmat et al. [7] proposed a technique for recognition of images using a convolutional neural network known as Alexnet. Aside from Alexnet CNN, the technique of Transfer Learning is employed, wherein a custom dataset specific to Malaysian food is compiled and subsequently utilized for Transfer Learning. The computer software achieved a high accuracy of 91.43 % owing to the extensive network of the Alexnet. Naritomi et al. [8] proposed a technique for reconstructing the three-dimensional form of a dish (including the food and plate) and a plate (without the food) from a single two-dimensional photograph. This method also allows for a more precise estimation of the volume of the meal. This suggestion was made because of the prevalence of 2D-based picture recognition in current approaches for determining food calorie values. Nevertheless, due to the fact that real food exists in three dimensions, the precision of calorie calculation using approaches based on twodimensional representations is inherently limited. Kasyap et al. [9] proposed a method that utilizes a deep learning algorithm to accurately measure calorie intake. In this context, measurements are obtained from images of food using various objects such as fruits and vegetables. This measurement is obtained using a neural network. A Tensor flow-based technique is utilized to compute the calorie content of food using a Convolutional Neural Network. The input to this computational model is a food image. The food calorie value is determined by employing the proposed CNN model in conjunction with food object detection. The main parameter of the outcome is the estimation of volume error, while the secondary parameter is the estimation of calorie error. The inaccuracy estimation of the volume is steadily decreased by 20 %. This suggests that the proposed CNN model is achieving a higher level of accuracy compared to the previous model. Hu et al. [10] proposed developing image-based calorie estimation models that can precisely recognize the names of Chinese and Western items, provide their calorie consumption and recipe details, and ultimately offer meal plan recommendations for various demographic groups. In order to determine the type of dish, they employed the Single Shot MultiBox Detector (SSD) for instantaneous analysis of object detection and categorization. In addition, they utilized a computer tool called "labelImg" to manually assign the appropriate original dish titles to the plates.

III. METHODOLOGY

A. System Configuration

The system configuration comprises a meticulously curated combination of software tools and frameworks, specifically tailored to seamlessly execute the food recognition and calorie estimates project. The system is built utilizing sophisticated deep learning frameworks such as OpenCV, TensorFlow, MiDaS, and Keras, with Python serving as the primary programming language[11]. The design aims to prioritize flexibility and dependability. Anaconda Navigator is a graphical user interface (GUI) that facilitates the management of packages and environments without the requirement of manually entering conda commands in a terminal window[12]. The versatile interface enables users to customize and organize workflows in the fields of data science, scientific computing, computational journalism, and machine learning. A modular design encourages the addition of extensions to enhance and augment functionality. By harnessing the collective capabilities of these software components, the system is poised to achieve outstanding performance in food recognition and calorie estimation, hence driving progress in the domains of computer vision and deep learning.

B. Algorithms

1) Convolutional Neural Network (CNN)

CNNs excel in automated food recognition and calorie estimation due to their ability to analyze visual data. Their hierarchical structure enables the extraction of complex features from images, facilitating accurate classification of food items. Trained on large datasets, CNNs can recognize and categorize food items while estimating their calorie content, facilitating automated dietary analysis[13], [14].

ISSN (Online) 2394-2320
International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Volume 11 Issue 08 August 2024

2) Multiple Depth Estimation Accuracy with Single Network (MiDaS)

MiDaS is used in automated food recognition and calorie estimation for its accurate monocular depth estimation from single images. Its high accuracy and generalization, achieved through comprehensive training, eliminate the need for multiple views or additional sensors. By providing depth information, MiDaS enhances the analysis of food items in images, leading to more precise estimations of portion sizes and calorie content.

3) OpenCV

OpenCV's robust image processing capabilities, including feature detection and machine learning functions, make it essential for automated food recognition and calorie estimation. With broad community support and cross-platform compatibility, OpenCV streamlines image preprocessing, feature extraction, and integration with machine learning models for accurate calorie content estimation from food images[15].

C. Working

1) Data Collection and pre-processing:

Data collection is a crucial aspect of constructing a machine learning model. It involves obtaining relevant information pertaining to specific variables in order to analyze and generate beneficial results. Nevertheless, a portion of the data may be subject to noise, meaning it may include imprecise, incomplete, or erroneous values. Therefore, it is imperative to preprocess the data prior to analyzing it and drawing conclusions. Data pre-processing encompasses the tasks of data cleansing, data transformation, and data selection[16].

2) Extraction of Features:

The pre-trained CNN model is used to process each image in the dataset. This model is specifically engineered to comprehend and analyze the visual material depicted in photographs. Throughout this procedure, the CNN model examines the complex characteristics of every image, detecting patterns, forms, textures, and other visual indicators[17]. The outcome of this analysis is a collection of top-level characteristics that succinctly represent the fundamental substance of the image. These features function as a depiction of the visual data in the image, transforming it into a format that is more readily understandable by the following steps of the captioning process.

3) Training and Testing:

This module involves training the CNN model using the preprocessed and normalized data. The dataset is commonly divided into training and testing sets. The training set is utilized to optimize the model parameters using gradient descent optimization techniques such as stochastic gradient descent (SGD) or Adam[18]. Subsequently, the testing set is employed to check the proficiency of the trained model on unfamiliar data, evaluating measures such as accuracy, precision, and recall. In this case, 80% of the dataset is divided into training sets, while the remaining 20% is allocated as testing sets.

4) Food Segmentation:

Neural networks, specifically convolutional neural networks (CNNs), can be used to segment food products based on their distinct visual features. Methods such as semantic segmentation and instance segmentation are employed to categorize and outline food items, enabling accurate identification and subsequent analysis[19]. This method enables a wide range of applications to be realized. This module segments food from the image and draws contours of the food.

5) Volume Estimation:

This module focuses on determining the area of the identified contours and estimating the depth of the food using MiDaS. The food volume is thereafter estimated using the aforementioned variables.

D. Engineering Standards

- IEEE 1858-2023- Camera Phone Image Quality.
- IEEE P2830 - Standard for Technical Framework and Requirements of Shared Machine Learning.

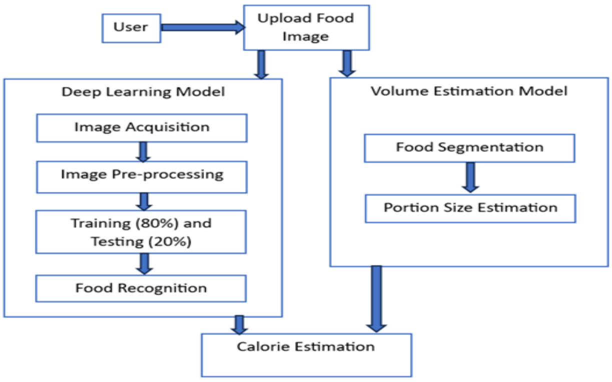

Figure 1: Block Diagram of the Automated Food Recognition and Calorie Estimation Model

Figure 2: The output of the Automated Food Recognition and Calorie Estimation model for a Parota

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 11 Issue 08 August 2024

IV. RESULTS AND DISCUSSION

A. Dataset Description

For this project, we have selected 15 food items and assigned labels to them. Each food undergoes training using a set of 30 distinct photos in order to determine the shared characteristics of the food. 80 % of the datasets are allocated for training, while the remaining 20 % are reserved for testing the CNN model.

Table 1: Dataset details

	Training	Testing
Dataset	80 %	20 %

B. Project outcome

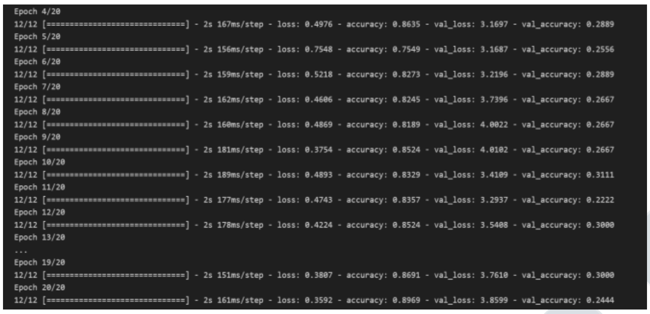

Figure 3: Accuracy of the manually trained model

The manually trained model performed well in our analysis, improving the accuracy of food recognition with each epoch. The manually trained model's baseline accuracy was 89 %.

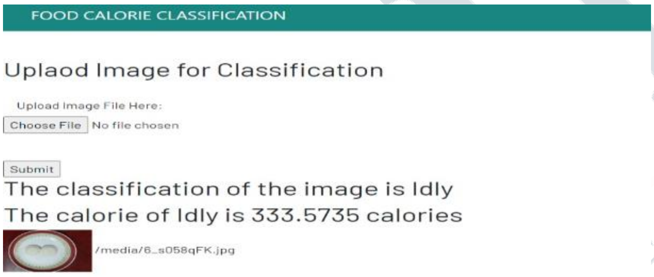

Figure 4: The output of the Automated Food Recognition and Calorie Estimation model for a plate of Idlies

Figure 5: The output of the Automated Food Recognition and Calorie Estimation model for a Burger

On the frontend, the user uploads the input image and then clicks the submit button. Next, the name of the food together with its calorie count is estimated and displayed beside the corresponding image.

The assessment outcomes of the Food Recognition and Calorie Estimation system demonstrate its robust performance in producing precise findings across a wide range of images. The qualitative assessments offer vital insights into the system's capacity to understand and precisely characterize the content of various images, hence strengthening its overall performance and effectiveness.

Furthermore, when contrasting the suggested model with the current models, significant disparities in performance indicators become apparent.

This indicates that our model not only succeeds in identifying the foods but also exhibits superior training performance, hence improving its capacity to make generalizations and function effectively on unfamiliar images.

Based on this thorough comparison, it is clear that the suggested model performs better than the existing model in terms of training performance. These findings emphasize the effectiveness and superiority of our model, showcasing its potential for diverse applications in the fields of computer vision and natural language processing.

V. CONCLUSION

Food recognition and calorie calculation using CNNs are a noteworthy advancement in the field of computational nutrition analysis. These systems have shown impressive ability in reliably recognizing different food items from photos and estimating their calorie content, providing significant information about dietary patterns and nutritional intake. By utilizing CNN algorithms, these systems have the capacity to transform the way we conduct dietary analysis, empowering individuals to make well-informed decisions regarding their nutrition and lifestyle.

Although significant advancements have been achieved, there remain unresolved difficulties and potential for additional improvement. Potential future advancements may prioritize enhancing the level of detail and precision in identifying food, integrating contextual data, and harnessing emerging technologies like multi-modal learning and reinforcement learning. Furthermore, it is imperative to engage in partnerships with specialists from many disciplines and prioritize ethical issues to guarantee the responsible and fair implementation of new technologies.

In general, the utilization of CNN algorithms for food recognition and calorie estimate shows great potential in encouraging healthy eating habits, aiding in weight management, and enhancing public health results. Ongoing research and advancements in this domain possess the capacity to enable individuals to exert authority over their food decisions, resulting in improved health and well-being for individuals and communities globally. The future of computational nutrition analysis seems promising, thanks to continuous developments and interdisciplinary collaboration. This field offers new opportunities for individualized and efficient dietary interventions.

VI. FUTURE PROSPECTS

Future advancements in food recognition and calorie estimate strive to better the accuracy, efficiency, and practicality of these technologies, fundamentally transforming the way we monitor and control our eating habits. Here is an examination of possible progressions in this domain:

- Future advancements in machine learning will mostly revolve around the enhancement of accuracy in food recognition through the continuous development of deep learning models. Advanced topologies, such as transformer-based networks and more sophisticated CNNs, are expected to offer improved accuracy and faster processing.
- Enhanced Training Datasets: By incorporating a greater variety of training data, food recognition models will acquire the capacity to accurately recognize a broader spectrum of meals, encompassing regional cuisines and specialized dietary items. Open-source endeavours and cooperative data-sharing could contribute to the expansion of these datasets.
- Integration of Augmented Reality (AR): AR technology has the potential to offer a more intuitive interface for identifying food and estimating calorie content. Users may effortlessly direct their smartphone or AR glasses towards their food, obtaining immediate input regarding the food's classification, approximate calorie count, and many nutritional details.
- Real-time processing of food recognition and calorie estimate will become more practical due to advancements in hardware, such as edge devices and GPUs. This would facilitate the use of mobile applications for monitoring and tracking fitness and health.
- Multimodal Inputs: The integration of visual, aural, and text inputs has the potential to enhance the accuracy of food recognition. This approach utilizes many modes of analysis, including food imagery, accompanying recipes, and audio signals such as voice orders, to enhance the accuracy of estimation.

REFERENCES

[1] M. Li, Y. Jiang, Y. Zhang, and H. Zhu, "Medical image analysis using deep learning algorithms," *Front. Public Heal.*, vol. 11, p. 1273253, Nov. 2023, doi: 10.3389/FPUBH.2023.1273253/BIBTEX.

[2] A. Abeltino *et al.*, "Digital applications for diet monitoring, planning, and precision nutrition for citizens and professionals: a state of the art," *Nutr. Rev.*, May 2024, doi: 10.1093/NUTRIT/NUAE035.

[3] M. Boujarra, A. Lechhab, A. Al Karkouri, I. Zrigui, Y. Fakhri, and S. Bourekkadi, "REVOLUTIONIZING LOGISTICS THROUGH DEEP LEARNING: INNOVATIVE SOLUTIONS TO OPTIMIZE DATA SECURITY," *J. Theor. Appl. Inf. Technol.*, vol. 29, no. 4, 2024, Accessed: May 24, 2024. [Online]. Available: www.jatit.org

[4] W. Tay, B. Kaur, R. Quek, J. Lim, and C. J. Henry, "Current Developments in Digital Quantitative Volume Estimation for the Optimisation of Dietary Assessment," *Nutr. 2020, Vol. 12, Page 1167*, vol. 12, no. 4, p. 1167, Apr. 2020, doi: 10.3390/NU12041167.

[5] M. M. Islam, "Unveiling the Power of Deep Learning: Insights into Advanced Neural n Network Architectures," *J. Artif. Intell. Gen. Sci. ISSN3006-4023*, vol. 3, no. 1, pp. 1–14, Feb. 2024, doi: 10.60087/JAIGS.V3I1.60.

[6] Y. Lu, "Food Image Recognition by Using Convolutional Neural Networks (CNNs)," *arXiv.org*, 2016.

[7] N. O. M. Salim, S. R. M. Zeebaree, M. A. M. Sadeeq, A. H. Radie, H. M. Shukur, and Z. N. Rashid, "Study for Food Recognition System Using Deep Learning," *J. Phys. Conf. Ser.*, vol. 1963, no. 1, Jul. 2021, doi: 10.1088/1742-6596/1963/1/012014.

[8] H. Kagaya, K. Aizawa, and M. Ogawa, "Food detection and recognition using convolutional neural network," *MM 2014 - Proc. 2014 ACM Conf. Multimed.*, pp. 1085–1088, Nov. 2014, doi: 10.1145/2647868.2654970.

[9] C. Liu, Y. Cao, Y. Luo, G. Chen, V. Vokkarane, and Y. Ma, "Deepfood: Deep learning-based food image recognition for computer-aided dietary assessment," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9677, pp. 37–48, 2016, doi: 10.1007/978-3-319-39601-9_4/TABLES/6.

[10] Y. Liang and J. Li, "Computer vision-based food calorie estimation: dataset, method, and experiment," May 2017, Accessed: May 24, 2024. [Online]. Available: https://arxiv.org/abs/1705.07632v3

[11] S. Sathyan and J. S. N., "Traffic Flow Prediction using Machine Learning Techniques - A Systematic Literature Review," *Int. J. Appl. Eng. Manag. Lett.*, vol. 6, no. 1, pp. 210–230, Apr. 2022, doi: 10.47992/IJAEML.2581.7000.0132.

[12] R. Mohanraj, R. Kailasam, V. P. Prakadeesh, R. Arun Prakash, B. Dhanushree, and M. Hasina, "IoT-Enabled Monitoring and Controlling System for Robotic Operations," *AIP Conf. Proc.*, vol. 2946, no. 1, Nov. 2023, doi: 10.1063/5.0178162/2920600.

[13] A. K. Sahoo and N. Mishra, "The Development of CNN-Based Food Scanning Model for Calorie Analysis: A Tool for Promoting Healthier Dietary Choices," *Proc. 2nd Int. Conf. Adv. Smart, Secur. Intell. Comput. ASSIC 2024*, 2024, doi: 10.1109/ASSIC60049.2024.10507809.

[14] G. A. Tahir and C. K. Loo, "A Comprehensive Survey of Image-Based Food Recognition and Volume Estimation Methods for Dietary Assessment," *Healthc. 2021, Vol. 9, Page 1676*, vol. 9, no. 12, p. 1676, Dec. 2021, doi: 10.3390/HEALTHCARE9121676.

[15] M. Mansouri, S. Benabdellah Chaouni, S. Jai Andaloussi, and O. Ouchetto, "Deep Learning for Food Image Recognition and Nutrition Analysis Towards Chronic Diseases Monitoring: A Systematic Review," *SN Comput. Sci.*, vol. 4, no. 5, pp. 1–17, Sep. 2023, doi: 10.1007/S42979-023-01972-1/TABLES/8.

[16] M. Kang and J. Tian, "Machine Learning: Data Pre-processing," *Progn. Heal. Manag. Electron.*, pp. 111–130, Sep. 2018, doi: 10.1002/9781119515326.CH5.

[17] Q. Han, Q. Yin, X. Zheng, and Z. Chen, "Remote sensing image building detection method based on Mask R-CNN,"

Complex Intell. Syst., vol. 8, no. 3, pp. 1847–1855, Jun. 2022, doi: 10.1007/S40747-021-00322-Z/FIGURES/5.

[18] E. M. Dogo, O. J. Afolabi, N. I. Nwulu, B. Twala, and C. O. Aigbavboa, "A Comparative Analysis of Gradient Descent-Based Optimization Algorithms on Convolutional Neural Networks," *Proc. Int. Conf. Comput. Tech. Electron. Mech. Syst. CTEMS 2018*, pp. 92–99, Dec. 2018, doi: 10.1109/CTEMS.2018.8769211.

[19] H. J. He, C. Zheng, and D. W. Sun, "Image Segmentation Techniques," *Comput. Vis. Technol. Food Qual. Eval. Second Ed.*, pp. 45–63, Jan. 2016, doi: 10.1016/B978-0-12-802232-0.00002-5.