Journal of Food and Nutrition Research (ISSN 1336-8672) Vol. *64*, 2025, No. 1, pp. 61–71

Deep learning algorithm for dessert recognition and nutritional evaluation

SEVILAY YUSUN – ONUR YILMAZ – FATIH TARLAK

Summary

Food recognition systems are crucial for healthcare and the food industry, aiding in diet tracking, personalised meal planning, and promoting nutritional awareness. This work develops a software interface that recognises food products using deep learning algorithms, and announces their nutritional values and gastronomic characteristics. Specifically, photographs of various desserts were captured in a restaurant setting, and the classification performance of two deep learning models, GoogleNet and ResNet-50, was analysed. Both models achieved high accuracy rates exceeding 99.6 %, with ResNet-50 demonstrating superior performance due to its lower error rates, higher accuracy, and faster learning capabilities. Based on these results, the interface was developed using ResNet-50 to provide consumers with detailed gastronomic information about desserts and support healthier dessert choices. At present, the resulting software is limited to the 23 dessert items on the menu of Healin restaurant (Nisantasi, Istanbul, Turkey), and the way they look in that particular restaurant, but the scope could be expanded in the future. This innovative approach enhances consumer awareness about healthy eating while offering a competitive edge for the food industry by effectively meeting consumer expectations.

Keywords

deep learning algorithms; food recognition; nutritional evaluation; dessert classification

Maintaining a healthy diet is essential for preventing various chronic diseases, such as heart disease, diabetes, and cancer, as highlighted by RUTHSATZ and CANDEIAS [1] and KAUR et al. [2]. Worldwide, there is increasing recognition of the significance of healthy eating habits, prompting many individuals to track their calorie intake and nutritional information to support or enhance their health [3, 4]. The ability to accurately predict and track calorie and nutrient intake is essential, especially in weight loss interventions and sustaining a healthy lifestyle [5]. A thorough understanding of food attributes, including their nutritional and caloric content, is fundamental for making informed dietary choices. Moreover, it is crucial to ensure the quality and safety of the food consumed [6–8].

Modern advancements in technology, such as electronic noses, computer vision, and spectroscopy, have become pivotal in detecting and analysing various food attributes [9–11]. These techniques offer sophisticated methods for identifying and quantifying different characteristics of food, enhancing our ability to ensure food quality and safety. However, despite these technological strides, there is still a significant demand for a method that is both rapid and precise for assessing food quality in everyday situations.

Pattern recognition and image processing have played a pivotal role in enhancing food intake reporting systems, particularly through advancements in convolutional neural networks (CNNs) and deep learning. These technologies enable more accurate food recognition, classification, and dietary assessment, providing valuable tools for individuals and professionals to monitor and analyse dietary intake effectively [12–15]. However, the task of accurately classifying food images presents

---

Sevilay Yusun, Department of Gastronomy and Culinary Arts, Faculty of Architecture and Design, Istanbul Gedik University, Cumhuriyet Street 1, 34876 Kartal, Istanbul, Turkey.
Onur Yılmaz, Department of Software Engineering, Faculty of Engineering, Istanbul Gedik University, Cumhuriyet Street 1, 34876 Kartal, Istanbul, Turkey.
Fatih Tarlak, Department of Bioengineering, Faculty of Engineering, Gebze Technical University, Cumhuriyet Street 2254, 41400 Gebze, Kocaeli, Turkey.

Correspondence author:
Fatih Tarlak, email: ftarlak@gtu.edu.tr

ongoing challenges. These challenges stem from various factors, including the wide diversity of food items, variations in presentation, and the complexity of distinguishing between visually similar foods [16, 17]. As a result, while current technologies have made progress, the need for more effective and efficient methods remains critical.

Analysing food products through images involves a comprehensive process that can be broken down into four key phases: food detection, classification, weight determination, and nutritional analysis. The advent of CNNs has significantly enhanced the accuracy of image identification, leading to a surge of interest in their application for food product analysis. CNNs have consistently outperformed traditional machine learning methods, a fact underscored by numerous studies that have modified models such as AlexNet to achieve higher accuracy in food recognition across various datasets [18–21]. These studies underscore the superior predictive capabilities of CNNs when compared to older, more conventional techniques [22, 23]. As a result, CNNs have become the preferred method for image-based food product analysis, offering enhanced precision and reliability in tasks that are critical for food detection and classification.

EXARCHOU et al. [24] focused on enhancing the design of a novel system intended to assist patients dealing with obesity or diabetes. Their approach involved using transfer learning and fine-tuning techniques to automatically differentiate between images of sweet and non-sweet foods by leveraging pre-trained CNN models. To further refine these deep neural networks, the researchers developed a new dataset derived from the original Food-101 dataset. This newly created dataset consisted of 19 000 images, encompassing both dessert and non-dessert foods and included 19 distinct categories of desserts. Among the models tested, the Google InceptionV3 architecture achieved the highest performance, with a validation accuracy of 95.9 %. To further validate the platform's practicality, the team also compiled an additional dataset consisting of food images captured under challenging conditions, such as varying lighting and shooting angles. This additional dataset was used to demonstrate the robustness and applicability of the system in real-world scenarios.

SATHISH et al. [25] focused on the development and analysis of an OpenCV-based deep learning model designed to identify Indian cuisine and accurately estimate its energy content. The researchers created an Indian Cuisine dataset by sourcing images from the internet and subjecting them to rigorous quality preprocessing. The classification of Indian dishes within this dataset was carried out using CNN, while the energy content of the identified foods was calculated using advanced image processing techniques. The results demonstrated that the developed model was highly effective, achieving an accuracy of 99.2 % on training data and 95.3 % on test data in detecting Indian cuisine dishes. Additionally, the model's energy estimation was impressively accurate, with an error margin of only ± 42 J (± 10 calories) when compared to the actual food. Despite these advancements, it is important to note that, to the best of our knowledge, there is no existing work specifically focused on the nutritional values and gastronomic evaluation of desserts using a developed software platform.

The aim of this study is to develop software that can accurately recognise food products and predict their nutritional content using deep learning algorithms. Specifically, the software is designed to evaluate the gastronomic characteristics of dessert products and provide consumers with valuable information to assist in managing their dietary choices, particularly when selecting desserts in restaurants. Additionally, the software offers restaurants and food manufacturers a tool for rapid dessert evaluation, leveraging fast data processing capabilities to enable the development of innovative dessert products. This approach represents a significant step toward gaining a competitive advantage and more effectively meeting consumer expectations.

MATERIALS AND METHODS

This work involves five critical stages, as illustrated in Fig. 1:

- Gathering real dessert images – collecting images of desserts before they were served at the restaurant.
- Compiling nutritional data and gastronomic details – assembling relevant nutritional information and gastronomic characteristics for the desserts.
- Employing deep learning algorithms – training deep learning algorithms using the collected dessert images.
- Creating software – developing the software using Matlab 9.10.0.1710957 (R2021a) software (MathWorks, Natick, Massachusetts, USA) to implement the trained algorithms.
- Evaluating and validating outcomes – assessing and validating the performance and accuracy of the developed software.

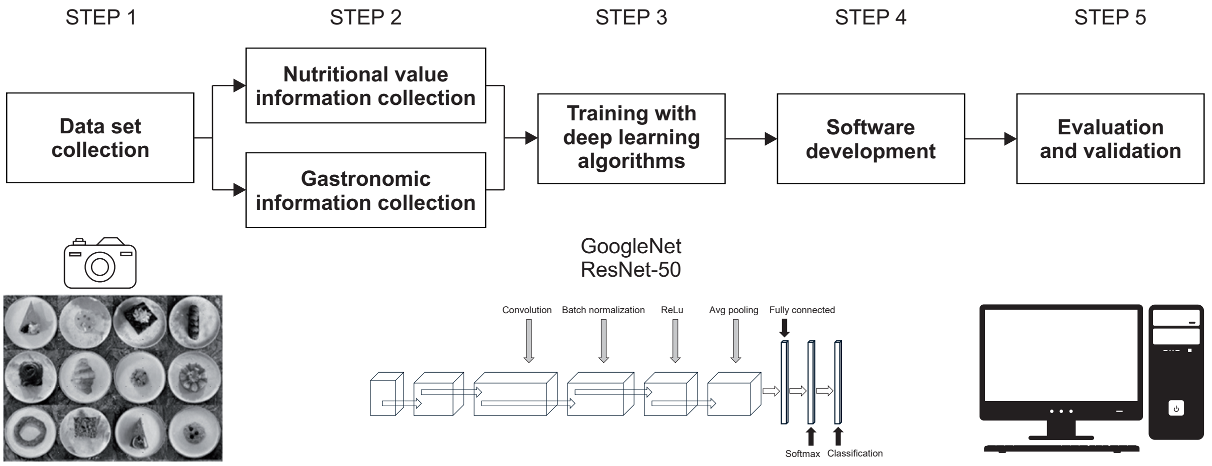

Fig. 1. Flow chart of the steps followed in the current study.

Comprehensive details for each of these steps were provided in the following subsections.

Dessert images

During the initial phase of this research, the focus was directed toward the conceptualisation and pre-training stages of the deep learning model. The methodology involved capturing high-resolution images of desserts featured on the menu of Healin restaurant (Istanbul, Turkey). Importantly, the dataset was composed of a consistent set of 23 dessert classes, which remained unchanged despite seasonal variations. The distribution of photographs for each specific dessert class is detailed in Tab. 1, offering a comprehensive overview of the dataset's composition and distribution across different classes. This detailed breakdown ensures a clear understanding of the dataset's structure, which is critical for the subsequent training and analysis phases of the model.

In the process of software development, genuine images representing all 23 dessert classes were meticulously gathered, resulting in an extensive collection totalling 10 267 images. To ensure a robust evaluation of the software, the team randomly selected a subset comprising 20 % of the total images, which were allocated for testing and validation purposes. This subset, consisting of 2 070 images, was specifically used to assess the software's performance, accuracy, and reliability. A sample of the dessert images utilised in this study is depicted in Fig. 2, providing a visual representation of the types of images included in the dataset.

Nutrition facts and gastronomic information

The Nutrition Information System (BeBIS, Istanbul, Turkey) is a highly advanced software application designed to provide access to national food databases from various countries. With a vast repository containing over 20 000 food items and comprehensive analysis covering more than 130 nutrients, BeBIS is an essential tool for evaluating the nutritional content of recipes, menus, and diets. Specifically in Turkey, BeBIS in-

Tab. 1. Number of photos of desserts in the dataset.

Class	Dessert	Number of photos
a	Blackberry cheesecake	711
b	Blue magic cheesecake	370
c	Brownie	550
d	Cherry chocolate tart	423
e	Chia pudding	366
f	Cinnamon roll	353
g	Croissant	425
h	Dark chocolate cookie	281
i	Fruit tart	600
j	Ganache tart	424
k	Gluten-free bagel	316
l	Granola bar	446
m	Huckleberry cheesecake	600
n	Ketogenic aronia cookies	282
o	Ketogenic cheesecake	461
p	Lemon matcha cheesecake	456
r	Mascarpone tart	593
s	Naked sponge cake	435
t	Oatmeal vegan cookie	314
u	Peanut butter energy ball	343
v	Snickers	505
y	Spanish cake	697
z	White chocolate chip cookie	316
Total		10 267

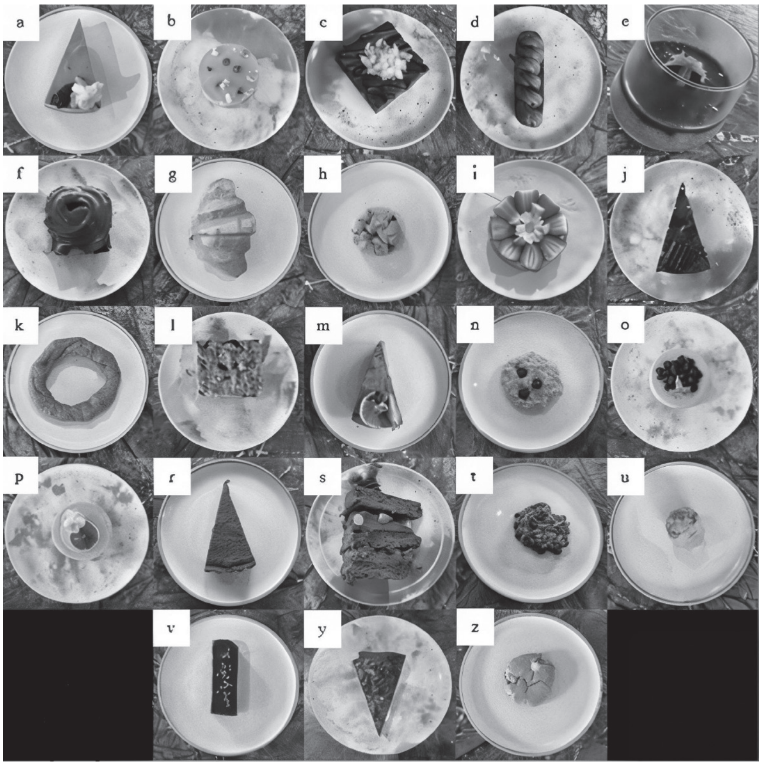

Fig. 2. Sample dessert images used to develop prediction software.

a – blackberry cheesecake, b – blue magic cheesecake, c – brownie, d – cherry chocolate tart, e – chia pudding, f – cinnamon roll, g – croissant, h – dark chocolate cookies, i – fruit tart, j – ganache tart, k – gluten-free bagel, l – granola bar, m – huckleberry cheesecake, n – ketogenic aronia cookies, o – ketogenic cheesecake, p – lemon matcha cheesecake, r – mascarpone tart, s – naked sponge cake, t – oatmeal vegan cookie, u – peanut butter energy ball, v – Snickers, y – Spanish cake, z – white chocolate chip cookie.

tegrates with the 50 national food composition databases (Turkomp) to deliver detailed nutritional information. This includes data on energy/calories, proteins, carbohydrates, vitamins, fats, antioxidant capacity, oxygen radical absorbance capacity value, amino acids, minerals, glycemic index, fatty acid compositions, and their respective percentage ratios, all based on the portion sizes entered by the user. This functionality makes BeBIS a powerful resource for dietitians, nutritionists, and anyone looking to accurately assess the nutritional quality of food.

Deep learning algorithms

Deep learning is a subset of machine learning that employs artificial neural networks to enable computers to learn from data and execute tasks by extracting meaningful features [26]. These neural networks are composed of multiple interconnected processing layers that work together to perform complex calculations in parallel, mimicking the structure and function of the biological neural system [27]. When trained on large datasets, deep learning models can achieve remarkable levels of accuracy in tasks such as object recognition, often surpassing human performance.

This study employed two state-of-the-art deep learning algorithms, GoogleNet and ResNet-50, from the Matlab Deep Learning Toolbox, to train a neural network for dessert recognition tasks. A dataset of 9 000 images was collected, representing 23 dessert categories, with roughly 390 images per category. The images were preprocessed to ensure uniform resolution and quality before training. To evaluate the system, an independent validation dataset consisting of 2 070 images (90 images per category) was used, ensuring no overlap with the training data.

GoogleNet, with 22 layers, and ResNet-50, with 50 layers, were selected for their advanced architectures tailored to image recognition tasks. GoogleNet employs an Inception framework to optimise computational efficiency while maintaining high accuracy. ResNet-50 uses a residual learning framework, enabling deeper networks by addressing the vanishing gradient problem. The training was conducted using the Stochastic Gradient Descent with Momentum (SGDM) algorithm. Both networks were trained over 50 epochs to achieve optimal performance. During training, features such as texture, colour, and patterns were progressively extracted and refined across multiple convolutional and pooling layers. Fig. 3 provides an overview of the architectures, showcasing the hierarchical process by which these models analyse and classify the input images to achieve high accuracy rates.

Software development

The main objective of this software development was to accurately identify dessert images from the menu of a restaurant and provide their nutritional and gastronomic details. The interface, built using the .NET Core framework (.NET Foundation, Redmond, Washington, USA), is designed for compatibility with both computers and smartphones, ensuring a user-friendly experience. The software is available for download [28], and mobile users can conveniently access it by scanning a QR code, simplifying the process of retrieving comprehensive nutritional and gastronomic information about the desserts offered at the restaurant.

Evaluation of training and validation process

To train the recognition software, a total of 9 000 food images were collected, with each of the twenty-three food categories represented by 90 images. To validate the software's performance, an additional set of twenty food product images per category was gathered, distinct from those used during the training phase. This resulted in a total of 2 070 images used specifically for validation purposes.

In classification problems, the performance of a classifier is typically evaluated using a confusion matrix, which provides a detailed breakdown of the classifier's predictions versus the actual outcomes. Based on the values within the confusion matrix, several important performance metrics can be calculated, including average accuracy (A_a), error rate (E_r), precision (PPV), recall (TPR), and F-score (F_{score}), by employing Eqs. 1–5, respectively, which are derived from the values within the matrix, as outlined by Sokolova and Lapalme [29]. These metrics offer a comprehensive understanding of the classifier's effectiveness and are crucial for assessing its overall accuracy and reliability:

$$A_a = \frac{\left(\sum_{i=1}^{l} \frac{tp_i + tn_i}{tp_i + fn_i + fp_i + tn_i}\right)}{l} \tag{1}$$

$$E_r = \frac{\left(\sum_{i=1}^{l} \frac{tp_i + fn_i}{tp_i + fn_i + fp_i + tn_i}\right)}{l} \tag{2}$$

$$PPV = \frac{\left(\sum_{i=1}^{l} \frac{tp_i}{tp_i + fp_i}\right)}{l} \tag{3}$$

$$TPR = \frac{\left(\sum_{i=1}^{l} \frac{tp_i}{tp_i + fn_i}\right)}{l} \tag{4}$$

$$F_{score} = \sum_{i=1}^{l} \frac{2 \times PPV \times TPR}{PPV + TPR} \tag{5}$$

where A_a refers to the average per-class effectiveness of a classifier, E_r refers to the average per-class classification error, PPV refers to the effectiveness of a classifier in identifying positive labels, TPR refers to the average per-class effectiveness of a classifier in identifying class labels, F_{score} refers to the relations between data's positive labels and those given by a classifier based on a per-class average, tp_i refers to the number true positive class, tn_i refers to the number true negative class, fp_i refers to the number false positive class, fn_i refers to the number false negative class, l is number of evaluated classes.

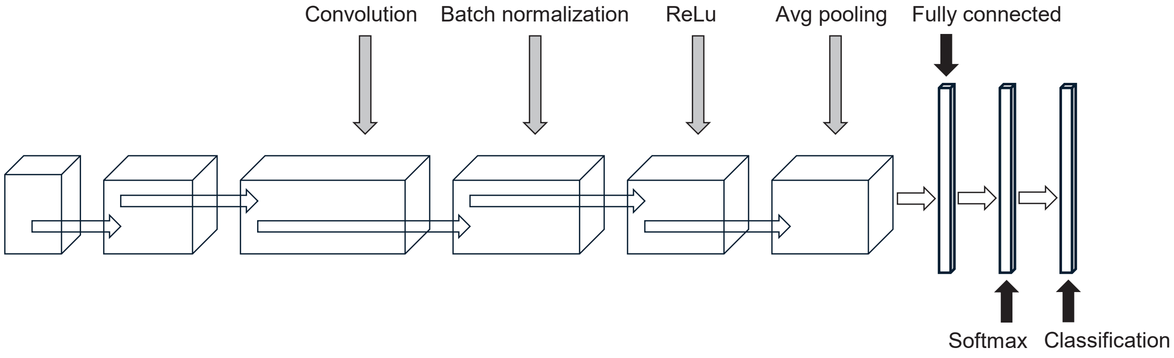

Fig. 3. Main components of used deep learning structures.

RESULTS AND DISCUSSION

In this study, which aimed to develop software for recognising food products using deep learning algorithms and pairing up this information with their nutritional content, the focus was placed on 23 dessert items featured on the fixed menu of the restaurant, which remained unchanged throughout the year, regardless of the season. The nutritional content of the recipes for these desserts was obtained through the BeBIS program, a comprehensive nutrition information system. The nutritional values were calculated per 100 g reference amount of the dessert. Detailed nutritional information for each dessert is provided in supplementary data (Tab. S1).

Among the desserts listed in Tab. S1, the three desserts with the highest energy values per 100 g of dessert are the ganache tart with 2 291 kJ, the brownie with 2 238 kJ, and the white chocolate cookie with 2 084 kJ. In contrast, the three desserts with the lowest energy values are the chia pudding with 267 kJ, the ketogenic cheesecake and the gluten free bagels with 962 kJ.

In terms of protein content per 100 g of dessert, the top three desserts are the peanut butter energy ball with 18 g of protein, the brownie with 15 g, and the ketogenic aronia cookie with 13 g. On the other hand, the desserts with the least protein content are the chia pudding with 2 g, the naked sponge cake with 4 g, and the cinnamon roll with 5 g.

The food images from the restaurant's menu were trained using the GoogleNet and ResNet-50 deep learning algorithms within Matlab software. GoogleNet utilised 22 layers, while ResNet-50 employed 50 layers for the training process. The dataset consisted of 10 267 dessert images, covering at least 282 unique images within each dessert category. This dataset was randomly divided, with 80 % allocated for training and the remaining 20 % reserved for testing purposes.

In the context of machine learning, particularly within statistical classification, a confusion matrix (also referred to as an error matrix) provides a visual representation of an algorithm's performance, especially in supervised learning. The matrix is structured with "actual" classes along the rows and "predicted" classes along the columns, allowing for a clear organisation of how actual class instances are classified versus how they are predicted by the model.

In this study, the confusion matrix was crucial in evaluating the test performance of both GoogleNet and ResNet-50. As illustrated in Fig. 4 and Fig. 5, the matrix distinguished between correct predictions and errors across various classes, corresponding to the dessert products shown in Fig. 2. During the testing phase, GoogleNet made 8 prediction errors, whereas ResNet-50 had only 4 false predictions out of a total of 2 070 samples (90 images per each of the 23 dessert classes). These results highlight ResNet-50's superior performance compared to GoogleNet during testing, as depicted in Fig. 4 and Fig. 5.

In this study, Tab. 2 presents the statistical evaluation metrics used, including A_a, E_r, PPV, TPR and F_{score}. The results indicate that GoogleNet achieved an A_a of 99.6 %, while ResNet-50 reached 99.8 %. These findings suggest that ResNet-50 outperformed GoogleNet during testing. The additional statistical metrics (E_r, PPV, TPR and F_{score}) further support this conclusion, consistently demonstrating that ResNet-50 exceeded GoogleNet in both accuracy and training efficiency.

Tab. 3 highlights the significant advancements of this study compared to previous research in food recognition systems. Earlier studies often used smaller datasets with fewer categories. In contrast, this study employed a dataset of 23 dessert categories, substantially increasing complexity while maintaining exceptional accuracy rates exceeding 99.6 % for both GoogleNet and ResNet-50. This superior performance, particularly by ResNet-50, reflects the robustness of the advanced CNN architectures used, which automatically extract hierarchical features and generalise well across diverse datasets without manual feature engineering. Such achievements underscore the study's contribution to handling more complex datasets while setting a new benchmark in food recognition accuracy.

The broader dataset and advanced methodologies make this work particularly relevant for practical applications, including dietary monitoring, restaurant menu digitisation, and food delivery systems. The demonstrated scalability and adaptability of the models position them as suitable solutions for diverse and complex food categories in global contexts. By addressing limitations in previous studies and achieving high performance with diverse data, this study establishes itself as a significant contribution to the field, paving the way for future research on larger datasets and optimisation for real-world deployment.

The training and testing results, along with the confusion matrices and statistical evaluation metrics, clearly establish ResNet-50 as the most efficient deep convolutional neural network in this study. However, it is important to note that

		Predicted class																						
		a	b	c	d	e	f	g	H	i	j	k	l	m	n	o	p	r	s	t	u	v	y	z
Actual class	a	90																						
	b		90																					
	c			89	1																			
	d				88																		2	
	e					90																		
	f						90																	
	g							90																
	h								90															
	i									90														
	j										90													
	k											90												
	l												90											
	m													90										
	n														90									
	o									2						85	3							
	p																90							
	r																	90						
	s																		90					
	t																			90				
	u																				90			
	v																					90		
	y																						90	
	z																							90

☐ correct prediction ■ error in prediction

Fig. 4. Confusion matrix for the testing process of GoogleNet.

a–z – class of dessert product (the marking as explained in Tab. 1).

		Predicted class																						
		a	b	c	d	e	f	g	H	i	j	k	l	m	n	o	p	r	s	t	u	v	y	z
Actual class	a	90																						
	b		90																					
	c			88							2													
	d				90																			
	e					90																		
	f						90																	
	g							90																
	h								90															
	i									90														
	j										90													
	k											90												
	l												90											
	m													90										
	n														90									
	o															90								
	p																90							
	r																	90						
	s																		90					
	t																			90				
	u																2				88			
	v																					90		
	y																						90	
	z																							90

☐ correct prediction ■ error in prediction

Fig. 5. Confusion matrix for the testing process of ResNet-50.

a–z – class of dessert product (the marking as explained in Tab. 1).

Tab. 2. Statistical evaluation for the testing process.

Network	A_a	E_r	PPV	TPR	F_{score}
GoogleNet	9.96×10^{-1}	3.86×10^{-3}	9.96×10^{-1}	9.96×10^{-1}	9.96×10^{-1}
ResNet-50	9.98×10^{-1}	1.93×10^{-3}	9.98×10^{-1}	9.98×10^{-1}	9.98×10^{-1}

A_a – average per-class effectiveness of a classifier, E_r – average per-class classification error, PPV – effectiveness of a classifier to identify positive labels, TPR – the average per-class effectiveness of a classifier to identify class labels, F_{score} – relations between data's positive labels and those given by a classifier based on a per-class average.

Tab. 3. Comparison of food recognition works published.

Source	Dataset	Category number	A_a [%]
Ege et al. [30]	Web image mining	15	80.6
Subhi and Ali [22]	Malaysian foods	11	74.7
Jeny et al. [31]	Bangladeshi foods	6	98.1
Razali et al. [32]	Sabahan foods	11	94.0
Shao et al. [4]	Junk food and fizzy drinks	14	80.1
Tarlak et al. [33]	Vegetables and fruits	45	99.8
Tarlak and Yucel [34]	Turkish breakfast	20	99.0

A_a – average per-class effectiveness of a classifier.

the training duration for deep learning algorithms can vary significantly. On an Intel(R) Core(TM) i5-1035G1 CPU @ 1.00 GHz processor (Intel, Santa Clara, California, USA), the first 100 iterations took 73 min for GoogleNet and 94 min for ResNet-50, indicating that GoogleNet had faster initial training speeds (Fig. 6).

In terms of file size, the deep learning toolkit in Matlab software shows that GoogleNet occupies 27.0 MB, while ResNet-50 requires 96 MB. This difference in file size reflects the varying complexity and scale of each network, with GoogleNet being more efficient in terms of training speed for less complex classification tasks.

However, when considering the overall learning times for all training images in this study, GoogleNet required 221 min, whereas ResNet-50 completed the training in 161 min. These comprehensive results firmly establish ResNet-50 as the superior deep learning algorithm, excelling not only in accuracy but also in processing speed, even when handling a full dataset of training images.

This study focuses on the recognition of food products using deep learning algorithms and the subsequent coupling with the information on their nutritional content. The performance of deep learning models, specifically GoogleNet and ResNet-50, was analysed using 23 different dessert items. The results demonstrated that both models achieved high accuracy rates, with ResNet-50 offering lower error rates and higher overall accuracy compared to GoogleNet.

To provide a more comprehensive analysis, the results indicate that ResNet-50 outperformed GoogleNet in terms of prediction accuracy, making fewer errors across the 23 dessert classes. This superior performance can be attributed to ResNet-50's deeper architecture, which employs residual connections to mitigate the vanishing gradient problem, enabling the model to learn more complex features. The implications of this finding are significant for practical applications in food recognition systems, as the improved accuracy of ResNet-50 suggests it could provide more reliable classifications in real-world scenarios, such as automated dietary monitoring or quality control in food production. However, it is also important to consider the computational cost associated with deeper architectures like ResNet-50, which

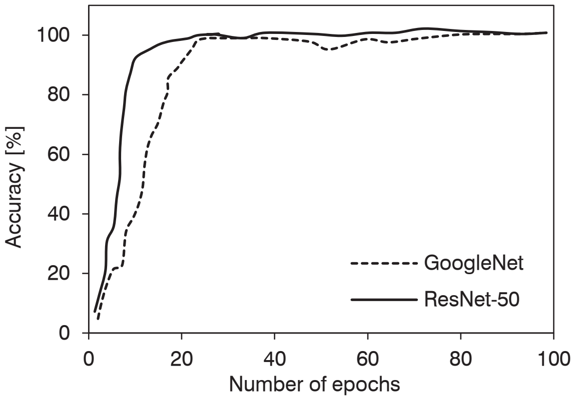

Fig. 6. Accuracy of the deep learning algorithms for training process by iteration.

may impact deployment on resource-constrained devices. Additionally, while GoogleNet performed slightly worse in this study, its lighter architecture could make it a more suitable choice in scenarios where computational efficiency is prioritised. Future research could address these trade-offs by exploring hybrid models or optimisation techniques to balance accuracy and computational efficiency. Moreover, investigating the types of errors made by each model and their underlying causes could offer valuable insights for further improving performance, particularly for challenging or visually similar classes.

When compared to other published studies, this research achieved notable success in food recognition, confirming the effectiveness of the deep learning models used. The findings indicate that deep learning techniques hold significant potential in the food industry for the identification and recognition of nutrient content. In particular, the ResNet-50 model was found to be more accurate and faster in tasks such as food recognition and nutrient content prediction. These results provide a strong foundation for future applications in the food industry, highlighting the practical value of deep learning in enhancing food product analysis and nutritional assessment.

CONCLUSION

This study successfully developed software for recognising dessert products using deep learning algorithms, and coupling the information with their nutritional values and gastronomic characteristics. The analysis focused on two prominent models, GoogleNet and ResNet-50, applied to 23 different dessert items from the restaurant. Both models achieved high accuracy rates exceeding 99 %, demonstrating their effectiveness in food recognition tasks. However, ResNet-50 stood out by offering lower error rates and faster learning capabilities, making it the more robust option for this application. The development of an interface based on ResNet-50 not only provides consumers with valuable gastronomic information but also supports informed decision-making in dessert choices. The findings of this study underscore the significant potential of deep learning approaches, particularly ResNet-50, in enhancing food recognition. This represents a notable advancement in the food industry, offering new avenues for improving food identification, quality control, and consumer education. Moreover, the results of this research set a solid foundation for future applications in this domain. The demonstrated accuracy and speed of ResNet-50 in handling complex classification tasks suggest that deep learning techniques can play a pivotal role in the continued evolution of the food industry, particularly in the areas of food safety, nutritional analysis, and personalised dietary recommendations. As such, this study contributes to the growing body of knowledge supporting the integration of advanced machine-learning models into practical, real-world applications within the food sector.

Acknowledgements

This study was financially supported by Istanbul Gedik University (Istanbul, Turkey) through the Center Supporting Scientific Research Projects. The authors would like to express their thanks to the academic staff working in the Nutrition and Dietetics Department at Istanbul Gedik University for helping collect the nutritional values of the desserts.

Supplementary data

Supplementary data related to this article can be found at https://www.vup.sk/en/download.php?bulID=2267.

REFERENCES

1. Ruthsatz, M. – Candeias, V.: Non-communicable disease prevention, nutrition and aging. Acta Biomedica, *91*, 2020, pp. 379–388. DOI: 10.23750/abm.v91i2.9721.
2. Kaur, R. – Kumar, R. – Gupta, M.: Deep neural network for food image classification and nutrient identification: A systematic review. Reviews in Endocrine and Metabolic Disorders, *24*, 2023, pp. 633–653. DOI: 10.1007/s11154-023-09795-4.
3. Soederberg Miller, L. M. – Cassady, D. L.: The effects of nutrition knowledge on food label use. A review of the literature. Appetite, *92*, 2015, pp. 207–216. DOI: 10.1016/j.appet.2015.05.029.
4. Nadeem, M. – Shen, H. – Choy, L. – Barakat, J. M. H.: Smart diet diary: Real-time mobile application for food recognition. Applied System Innovation, *6*, 2023, article 53. DOI: 10.3390/asi6020053.
5. Sawamoto, R. – Nozaki, T. – Nishihara, T. – Furukawa, T. – Hata, T. – Komaki, G. – Sudo, N.: Predictors of successful long-term weight loss maintenance: a two-year follow-up. BioPsychoSocial Medicine, *11*, 2017, article 14. DOI: 10.1186/s13030-017-0099-3.
6. Khan, M. A. – Khan, R. – Praveen, P. (Eds.).: Artificial intelligence applications in agriculture and food quality improvement. Hershey : IGI Global, 2022. ISBN: 9781668451410. DOI: 10.4018/978-1-6684-5141-0.
7. Khan, R.: Artificial intelligence and machine learning in food industries: a study. Journal of Food Chemistry and Nanotechnology, *7*, 2022, pp. 60–67. DOI: 10.17756/jfcn.2021-114.

8. Khan, R. – Kumar, S. – Dhingra, N. – Bhati, N.: The use of different image recognition techniques in food safety: a study. Journal of Food Quality, *2021*, 2021, article 7223164. DOI: 10.1155/2021/7223164.
9. Seesaard, T. – Goel, N. – Kumar, M. – Wongchoosuk, C.: Advances in gas sensors and electronic nose technologies for agricultural cycle applications. Computers and Electronics in Agriculture, *193*, 2022, article 106673. DOI: 10.1016/j.compag.2021.106673.
10. Tarlak, F. – Ozdemir, M. – Melikoglu, M.: Computer vision system approach in colour measurements of foods: Part I. Development of methodology. Food Science and Technology, *36*, 2016, pp. 382–388. DOI: 10.1590/1678-457X.11615.
11. Tarlak, F. – Ozdemir, M. – Melikoglu, M.: Computer vision system approach in colour measurements of foods: Part II. Validation of methodology with real foods. Food Science and Technology, *36*, 2016, pp. 499–504. DOI: 10.1590/1678-457X.02616.
12. Allegra, D. – Battiato, S. – Ortis, A. – Urso, S. – Polosa, R.: A review on food recognition technology for health applications. Health Psychology Research, *8*, 2020, article 9297, pp. 172–187. DOI: 10.4081/hpr.2020.9297.
13. Jiang, L. – Qiu, B. – Liu, X. – Huang, C. – Lin, K.: DeepFood: food image analysis and dietary assessment via deep model. IEEE Access, *8*, 2020, pp. 47477–47489. DOI: 10.1109/ACCESS.2020.2973625.
14. Jiang, S. – Min, W. – Liu, L. – Luo, Z.: Multi-scale multi-view deep feature aggregation for food recognition. IEEE Transactions on Image Processing, *29*, 2020, pp. 265–276. DOI: 10.1109/TIP.2019.2929447.
15. Lo, F. P. W. – Sun, Y. – Qiu, J. – Lo, B.: Image-based food classification and volume estimation for dietary assessment: a review. IEEE Journal of Biomedical and Health Informatics, *24*, 2020, pp. 1926–1939. DOI: 10.1109/JBHI.2020.2987943.
16. Yang, S. – Chen, M. – Pomerleau, D. – Sukthankar, R.: Food recognition using statistics of pairwise local features. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA. New York : IEEE, 2010, pp. 2249–2256. ISBN: 978-1-4244-6985-7. DOI: 10.1109/CVPR.2010.5539907.
17. Boushey, C. – Spoden, M. – Zhu, F. – Delp, E. – Kerr, D.: New mobile methods for dietary assessment: review of image-assisted and image-based dietary assessment methods. Proceedings of the Nutrition Society, *76*, 2017, pp. 283–294. DOI: 10.1017/S0029665116002913.
18. Bossard, L. – Guillaumin, M. – Van Gool, L.: Food-101 – Mining discriminative components with random forests. In: Fleet, D. – Pajdla, T. – Schiele, B. – Tuytelaars, T. (Eds.): Computer Vision–ECCV 2014. Lecture notes in computer science, vol. 8694. Cham : Springer, 2014, pp. 446–461. ISBN: 978-3-319-10598-7. DOI: 10.1007/978-3-319-10599-4_29.
19. Kawano, Y. – Yanai, K.: Food image recognition with deep convolutional features. In: UbiComp '14 Adjunct: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication. New York : Association for Computing Machinery, 2014, pp. 589–593. DOI: 10.1145/2638728.2641339.
20. Kawano, Y. – Yanai, K.: Automatic expansion of a food image dataset leveraging existing categories with domain adaptation. In: Agapito, L. – Bronstein, M. – Rother, C. (Eds.): Computer Vision – ECCV 2014 Workshops. ECCV 2014. Lecture notes in computer science, vol. 8927. Cham : Springer, 2015, pp. 3–17. ISBN: 978-3-319-16199-0. DOI: 10.1007/978-3-319-16199-0_1.
21. Krizhevsky, A. – Sutskever, I. – Hinton, G. E.: ImageNet classification with deep convolutional neural networks. Communications of the ACM, *60*, 2017, pp. 84–90. DOI: 10.1145/3065386.
22. Subhi, M. A. – Ali, S. M.: A deep convolutional neural network for food detection and recognition. 2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), Sarawak, Malaysia. New York : IEEE, 2018, pp. 284–287. DOI: 10.1109/IECBES.2018.8626720.
23. Shirmard, H. – Farahbakhsh, E. – Heidari, E. – Beiranvand Pour, A. – Pradhan, B. – Müller, D. – Chandra, R.: A comparative study of convolutional neural networks and conventional machine learning models for lithological mapping using remote sensing data. Remote Sensing, *14*, 2022, article 819. DOI: 10.3390/rs14040819.
24. Exarchou, D.-M. – Alexiadis, A. – Triantafyllidis, A. – Ioannidis, D. – Votis, K. – Tzovaras, D.: A deep learning-based dessert recognition system for automated dietary assessment. In: Gao, X. – Jamalipour, A. – Guo, L. (Eds): Wireless mobile communication and healthcare. MobiHealth 2021. Lecture notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol. 440. Cham : Springer, 2022, pp. 60–70. ISBN: 978-3-031-06367-1. DOI: 10.1007/978-3-031-06368-8_4.
25. Sathish, S. – Ashwin, S. – Quadir, M. A. – Pavithra, L. K.: Analysis of convolutional neural networks on Indian food detection and estimation of calories. Materials Today: Proceedings, *62*, 2022, pp. 4665–4670. DOI: https://doi.org/10.1016/j.matpr.2022.03.122.
26. Alzubi, J. – Nayyar, A. – Kumar, A.: Machine learning from theory to algorithms: an overview. Journal of Physics: Conference Series, *1142*, 2018, article 012012. DOI: 10.1088/1742-6596/1142/1/012012.
27. Prieto, A. – Prieto, B. – Ortigosa, E. M. – Ros, E. – Pelayo, F. – Ortega, J. – Rojas, I.: Neural networks: an overview of early research, current frameworks and new challenges. Neurocomputing, *214*, 2016, pp. 242–268. DOI: 10.1016/j.neucom.2016.06.014.
28. Yusun, S. – Yılmaz O. – Tarlak, F.: AI-based-dessert-recognition-and-nutritional-analysis. In: GitHub [online]. San Francisco : GitHub, uploaded 28 August 2024 [cited 28 August 2024]. <https://github.com/ftarlak/AI-based-dessert-recognition-and-nutritional-analysis>
29. Sokolova, M. – Lapalme, G.: A systematic analysis of performance measures for classification tasks. Information Processing and Management, *45*, 2009, pp. 427–437. DOI: 10.1016/j.ipm.2009.03.002.

30. Ege, T. – Yanai, K.: Simultaneous estimation of food categories and calories with multi-task CNN. 2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA), Nagoya, Japan. New York : IEEE, 2017, pp. 198–201. DOI: 10.23919/MVA.2017.7986835.
31. Jeny, A. A. – Junayed, M. S. – Ahmed, I. – Habib, M. T. – Rahman, M. R.: FoNet – local food recognition using deep residual neural networks. 2019 International Conference on Information Technology (ICIT), Bhubaneswar, India. New York : IEEE, 2019, pp. 184–189. DOI: 10.1109/ICIT48102.2019.00039.
32. Razali, M. N. – Moung, E. G. – Yahya, F. – Hou, C. J. – Hanapi, R. – Mohamed, R. – Hashem, I. A. T.: Indigenous food recognition model based on various convolutional neural network architectures for gastronomic tourism business analytics. Information, *12*, 2021, article 322. DOI: 10.3390/info12080322.
33. Tarlak, F. – Yücel, Ö. – Yılmaz, O.: Food recognition and nutrition facts determination with deep convolution neural network models. Carpathian Journal of Food Science and Technology, 2023, Special issue, pp. 1–10. DOI: 10.34302/SI/236.
34. Tarlak, F. – Yücel, Ö.: Identification of foods in the breakfast and determination of nutritional facts using deep convolution neural networks. Food Science and Technology, *43*, 2023, article e10323. DOI: 10.5327/fst.10323.

Received 14 August 2024; 1st revised 29 November 2024; accepted 13 January 2025; published online 14 March 2025.