Indonesian Journal of Electrical Engineering and Computer Science
Vol. 38, No. 2, May 2025, pp. 1331~1338
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v38.i2.pp1331-1338

Accurate segmentation of fruit based on deep learning

Esraa Abu Elsoud[1], Omar Alidmat[1], Suhaila Abuowaida[2], Esraa Alhenawi [1], Nawaf Alshdaifat[3], Ahmad Aburomman[1], Huah Yong Chan[4]
[1]Department of Computer Science, Faculty of Information Technology, Zarqa University, Zarqa, Jordan
[2]Department of Computer Science, Faculty of Information Technology, Al-Albayt University, Mafraq, Jordan
[3]Faculty of Information Technology, Applied Science Private University, Amman, Jordan
[4]School of Computer Sciences, Universiti Sains Malaysia, Gelugor, Malaysia

Article Info

Article history:

Received Aug 10, 2024
Revised Nov 10, 2024
Accepted Nov 24, 2024

Keywords:

Deep learning
Mask R-CNN
Multiple fruit
Region proposal network
Segmentation

ABSTRACT

In the last few years, deep learning has exhibited its efficacy and capacity in the field of computer vision owing to its exceptional precision and widespread acceptance. The primary objective of this study is to investigate an improved approach for segmentation in the context of various fruit categories. Despite the utilization of deep learning, the current segmentation techniques for various fruit items exhibit subpar performance. The proposed enhanced multiple fruit segmentation algorithm has the following main steps: 1) modifying the size of the filter, 2) the process of optimizing the ResNet-101 block involves selecting the most suitable count of repetitions. The multiple fruit dataset is split 80% in the training stage and 20% in the testing stage. These images were utilized to train a deep learning (DL) based algorithm, which aims to identify multiple fruit items within images accurately. The proposed algorithm has a lower training time compared to the other algorithms. The thresholds exhibit greater values compared to the thresholds of state-of-the-art algorithms.

Corresponding Author:

Suhaila Abuowaida
Department of Computer Science, Faculty of Information Technology, Al-Albayt University
Mafraq, Jordan
Email: suhila@aabu.edu.my

1. INTRODUCTION

Fruit image segmentation is a significant and complex problem within the field of computer vision [1]. This task is of importance as it directly impacts and facilitates dietary assessment, as fruit segmentation serves as a fundamental step in these processes. Moreover, the process of fruit segmentation presents several challenges, mostly due to the presence of multiple fruit items that have similar shapes and appearances. Consequently, this similarity hinders the accurate differentiation of fruit solely based on their visual representations. Furthermore, several fruit items exhibit a diverse range of shapes, colors, and sizes, hence posing a greater difficulty in the classification of fruit types. The process of separating irregularly shaped fruit items in a multiple-fruit image becomes challenging, particularly when there is an occlusion present. Moreover, the presence of diverse quantities of fruit depicted in the image is a significant determinant that influences the approach employed for fruit recognition. This factor is vital in the advancement of multiple fruit segmentation. Hence, the development of multiple fruit segmentation poses a significant challenge. Numerous algorithms have pro-posed distinct methodologies for addressing the task of fruit segmentation [2]. However, the majority of studies have focused on the identification of multiple fruit items based on a single image [3]. The segmentation of fruit items was not possible due to the challenges that are associated with this specific task.

The objective of the paper is to establish a theoretical build for the aim of fruit segmentation to solve the issues mentioned above. Hence, a special algorithm referred to as multiple fruit segmentation was introduced, comprising two stages. The first stage improves fruit feature extraction by improving RestNet-101, resulting in an improved ResNet-101 backbone to solve low detection rates. The next step involves the implementation of the region proposal network (RPN) to simplify the localization of different fruit items. The suggested algorithm focuses on the task of multiple fruit segmentation. The enhanced multiple fruit segmentation technique employs the region of interest (RoI), for segmentation, to extract features specific to each instance. The evaluation of the multiple fruit segmentation algorithm consists of two measurements: the average over union section over IoU (AP) with various thresholds and the duration of training and testing. The measurements were further compared with the algorithms Mask R-CNN [4], YOLACT [5], and CASCADE R-CNN [6]. The factors contributing to the variations observed in the outcomes produced by these four algorithmshave been identified. The paper concludes with a summary of the results.

2. RELATED WORK

This section discusses a comprehensive examination of the algorithms described in prior work, focusing on their use in identifying, detecting, and segmenting fruits. These algorithms encompass advanced deep-learning approaches. Deep learning has demonstrated its effectiveness in computer vision [7]-[11]. It has been used to detect fruit based on images, distinguished by its high accuracy, as presented in this section. In this respect, convolutional neural network (CNN) is considered a state-of-the-art algorithm in deep learning image segmentation. Furthermore, [12] used CNN to segment fruit image tasks. CNN demonstrated that it was more accurate and could manage large datasets than traditional machine learning algorithms. The algorithm based on deep learning was applied to segment fruit using combined selective search, CNN, and GrabCut algorithms [13]. The main issue in the algorithm was an overlapping error, especially in the edges. A new segmentation algorithm was presented: a fully automated novel fruit intake segmentation system for semantic segmentation using FCN [14]. The benefit of this algorithm is enhanced reliability in terms of segmentation types. On the other hand, this algorithm could not handle the same instance of the fruit segment.

A semantic segmentation algorithm was applied to find the portion of the fruit [15]-[17]. The algorithm combined FCN and fast R-CNN, which decreased the time speed in the fruit segmentation. The Mask R-CNN was used in fruit image segmentation [18]-[24]. The proposed algorithms showed encouraging improvements in the quality of fruit image segmentation than traditional machine learning algorithms. However, the algorithms relied on Mask R-CNN. This lost certain features at the instance segmentation level and required a lengthy training period. This study is concerned with the segmentation of fruit images. Various approximation algorithms have been investigated, and researchers usually work on the main problem: fruit image segmentation. In the literature, fruit image segmentation still includes several weaknesses, such as not being flexible when dealing with diverse shapes, colors, and sizes and difficulty in separating those types having irregular shapes, especially when there is an occlusion in the fruit image.

The instance segmentation for multiple foods suffers from poor performance despite the use of deep learning due to the adoption of ResNet-101 as a backbone for feature extraction. The ResNet-101 problems still exist, such as determining a suitable number of ResNet-101 blocks, and small features become smaller or vanish during downsampling. An enhanced instance segmentation algorithm for multiple types of fruit can be designed and enhanced. This proposed algorithm enhanced the ResNet-101 backbone for better feature extraction by finding the suitable number of ResNet-101 blocks and connecting them with an additional convolutional layer

3. THE DATASET

The successful execution of segmentation tasks utilizing deep learning algorithms requires access to large datasets and significant computational resources. The problem of computational powerhas been successfully solved by the utilization of graphics processing units (GPUs). However, there exists alimited number of databases that contain meticulously annotated open-source data, particularly in the domain of segmentation, with a specific focus on multiple fruit segmentation. As a result, a dataset was generated that includes several fruit items to help with the task of multiple fruit segmentation. In addition, the training algorithm included the utilization of the COCO dataset [25], [26], which both included food-related content. The images used in this research shows various levels of resolution.

Several challenges arise when selecting images that encompass a range of characteristics. The multiple fruit dataset contains 27 distinct classes. Hence, the dataset comprising several fruit images encompasses a total of 27,000 images. The multiple fruit dataset issplit 80% in the training stage and 20% in the testing stage.

4. METHOD

The paper provides significant insights into the segmentation of multi-food. It focuses on identifying tasks related to multiple fruit items, as well as localizing and segmenting each fruit item based on the input image. These tasks pose challenges due to various factors, such as the presence of multiple fruit items and the variations in color and size between them. Furthermore, this study introduces a novel algorithm that enhances many approaches, such as CNN, region proposal networks (RPN), and segmentation algorithms. The aim is to achieve optimal results by accurately representing fruit images.

4.1. The enhanced multiple fruit segmentation algorithm

A novel approach for enhanced multiple fruit segmentation has been developed, leveraging an improved version of the ResNet-101 architecture as its backbone. This modified deep learning model has been specifically optimized to extract more nuanced and comprehensive fruit features. The improved ResNet-101 backbone uses cutting-edge methods like dilated convolutions and attention mechanisms to gather both small details and big-picture information at the same time. This results in more accurate and robust fruit segmentation, even in complex scenarios with multiple overlapping fruits or varying lighting conditions. The improved algorithm not only excels in distinguishing between different types of fruits but also precisely delineates their boundaries, making it particularly valuable for applications in automated harvesting, quality control in the foodindustry, and dietary analysis systems.

4.1.1. Improving ResNet-101 backbone for better feature extraction

CNN is a type of deep learning architecture that relies on a variety of inputs, such as audio, picture, text, and video. The high accuracy of CNN has been shown in various domains. Several widely recognized CNN architectures that serve as backbone networks include AlexNet, VGG, and ResNet-101. The ResNet-101 architecture has been identified as a suitable approach for achieving improved accuracy in object detection [27]. The ResNet-101 model achieves high performance by employing a deep neural network architecture that incorporates a series of blocks designed to solve the problem of gradient vanishing. This is achieved by the addition of shortcut connections [27].

The architecture of ResNet-101 incorporates two types of shortcut connections. 1) Identity block: if the dimensions of the input and output are the same, then the identity block can be applied directly. 2) Convolutional block: if the size of the input differs from that of the output, a convolution block is used to adjust the size of the input. The gradient vanishing problem has been solved by the ResNet-101 backbone with the implementation of shortcut connections. However, ResNet-101 shows several issues: The utilization of a large filter size in the initial convolution layer results in the extraction of less complex features. ResNet-101 blocks might have received a limited amount of repetitions, while others may have been excessively repeated. There remains a necessity to improve the detection rate for fruit recognition.

The enhanced ResNet-101 architecture has three stages. 1) The first convolution layer should be modified to contain a smaller filter size. 2) The ResNet-101 block should be optimized by carefully selecting the appropriate number of repetitions. 3) To enhance the model's performance, a convolutional layer should be introduced at each level. The suggested enhancement to the ResNet101 backbone involves reducing the filter size in the initial convolution layer and arranging the layers that require more iterations. Additionally, a convolutional layer is added for each stage to enhance the feature selection process by extracting multiple specific and exact feature maps from the shallow layers. Figure 1 shows the improved ResNet-101 backbone. Thus, the present study has developed a novel approach for detecting several foods, showing reduced training and testing duration and enhanced accuracy, solving the above mentioned issues.

4.1.2. Filter size modification

The filters that are used by the CNN are utilized to find the feature map from the input image by using a suitable filter size. Various filter sizes are used to extract specific details from the images within the CNN layer. The goal of this study is to investigate the techniques utilized for fruit recognition. Hence, since many computations are required in this layer to produce a singular value as a feature map, it is advisable to avoid using 7×7, as they frequently simplify the features and may overlook important detailscontained in the images. On the other hand, choosing a smaller filter size, such as 3×3, provides a greater amount of detail, therefore leading to enhanced accuracy [28], [29]. Therefore, to extract a greater number of features, the suggested backbone employed a smaller filter size of 3×3 instead of the 7×7 filter size applied in ResNet-101.

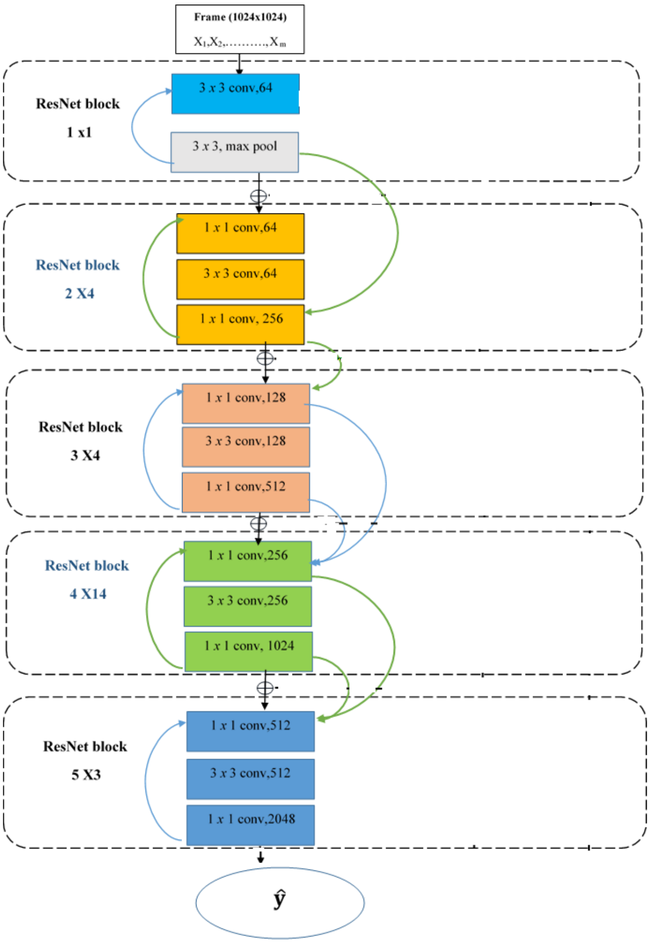

Figure 1. Enhanced ResNet-101 backbone

4.1.3. Optimizing the ResNet-101 block

Which prevents the achievement of improved accuracy and reduced training and testing times. The frequency of iterations is established through experimentation performed on the designed ResNet-101 backbone about fruit data. The experiment includes evaluating the accuracy of the ResNet block. If the new accuracy passes the old accuracy, the ResNet block is recursively repeated until a higher level of accuracy is achieved. The repetitions stop when the new accuracy comes under the old accuracy. As an outcome, the process of optimization involves determining the appropriate number of repeats. The aforementioned procedure has been completed to identify the blocks that require additional iterations or fewer iterations. This procedure is executed to enhance the computational time efficiency. After that, the results of each step are combined with an extra convolutional layer to make the suggested ResNet-101 backbone more accurate.

To localize and segment food, the enhanced multiple fruit segmentation algorithm incorporates the RPN and the mask branch suggested by citeli2017fully. This algorithm extracts features specific to each RoI where fruit is identified. The mask branch produces a mask with dimensions' m x m for every RoI according to [4]. The total number of classificationsin this context can be represented as K+1, where K denotes the count of detectable items present in the image.The additional class is specifically designated for the background. A CNN-based mask classifier is employed to classify the fixed features after their creation. Multiple masks are produced for different objects, with no competition observed between classes. Conventional neural networks typically employ the regions chosen by RoI classifiers as input to generate a mask. This ensures that spatial locations and features are suitably sized. The resulting matrix consists of elements with a value of 1 indicating the likelihood of an object's presence, while elements with a value of 0 indicate the absence of any items. This matrix is commonly referred to as a binary mask. The mask head utilizes the fully convolutional network (FCN) to predict the mask at a pixel level for each RoI. The present study employs the specialized classification branch proposed by [4] to make predictions regarding the identification of mask IDs. The output of the enhanced multiple fruit segmentation algorithm encompasses classification labels, bounding boxes, and masks for each identified fruit item, as illustrated in Figure 1. The loss function for this model comprises three components: classification loss (Lcls), bounding box loss (Lbox), and mask loss (Lmask). The mask loss is calculated as the average binary cross-entropy loss, as defined by [4]. The overall loss function for multiple fruit segmentation can be mathematically represented as (1).

$$L = Lcls + Lbox + Lmask \quad (1)$$

5. EXPERIMENTAL RESULTS AND RESULT

The experimental results are detailed in this section, and the evaluation of the algorithm is discussed in the preceding sections. The evaluation of the multiple fruit segmentation algorithm was conducted using two measurements: The study employed several thresholds of average precision (AP) to assess and compare the performance of multiple fruit segmentation algorithms. Additionally, the study measured the time required for both the training and testing phases.

5.1. Experimental parameters

An algorithm for multiple fruit segmentation was implemented using TensorFlow. The methodologies were assessed with Amazon Web Services (AWS) and Amazon Machine Image (AMI), employing the GPU-Us-Tesla V100 with 16 GB of memory and VCPUs-8 cores with 61 GB of memory. Furthermore, the optimization algorithm utilized in this work was stochastic gradient descent (SGD). The weight decay was established at 0.0002, with a learning momentum of 0.8 and a learning rate of 0.002, over the course of 40 epochs. The epoch consisted of 1000 iterations.

5.2. Optimization ResNet-101 backbone

The precision of fruit segmentation is augmented by incorporating supplementary convolutional layers at each tier and selecting an optimised backbone that accounts for the exact number of repetitions within each neural block. The optimisation procedure was executed using fruit photographs as the basis. The integration of convolutional layers at every level enhances the accuracy of fruit recognition. This technique involves gathering local feature maps from each stage, subsequently leading to the extraction of the final feature maps. The ultimate feature map is obtained by summing the local feature maps from each stage. Thereafter, the feature map functions as the input for the subsequent element in the multi-fruit recognition procedure. The RPN functions within the faster R-CNN architecture.

The efficacy of the augmented ResNet-101 model is illustrated in Figure 1. Experiments were performed for each level of ResNet-101, and the ideal number of repetitions of the ResNet-101 block was established post-training. Subsequently, supplementary convolutional layers were integrated into each level. The advised number of repetitions for the augmented ResNet-101 block, following the incorporation of the convolution layer, is detailed below: The proposed architecture has one convolution block succeeded by three repeating identity blocks. The proposed architecture has one convolution block succeeded by three repeating identity blocks. The proposed architecture has one convolution block succeeded by 13 repeating identity blocks. The proposed architecture has one convolution block succeeded by two repeated identity blocks.

5.3. Evaluation results of the enhanced multiple fruit segmentation algorithm with different state-of-the-art algorithms

A comparative analysis was carried out to evaluate the performance of the suggested algorithm in relation to the most advanced algorithms. Subsequently, these algorithms underwent training and testing using different criteria. The contrast is shown in Table 1. The result demonstrates that the suggested algorithm displays superior results. The suggested enhanced multiple fruit segmentation algorithm yields optimal results owing to the refined ResNet-101 backbone, achieved by selecting block duplicates and reducing filter size. Additionally, by incorporating the convolutional layer. The efficacy of the improved multiple fruit segmentation algorithm in comparison to other leading algorithms. The results indicate that the suggested enhanced multiple fruit segmentation algorithm offers a superior solution. The proposed approach enhances various fruit examples by utilizing the visual testing results from other algorithms, as illustrated in the Figure 2.

Table 1. Evaluation results of the enhanced multiple fruit segmentation algorithm with different algorithms

Algorithms	AP50	AP75	AP90
Mask R-CNN	91.02	77.82	68.97
YOLACT	89.27	77.48	68.83
CASCADE R-CNN	94.41	82.71	71.02
Proposed algorithm	97.13	86.22	72.99

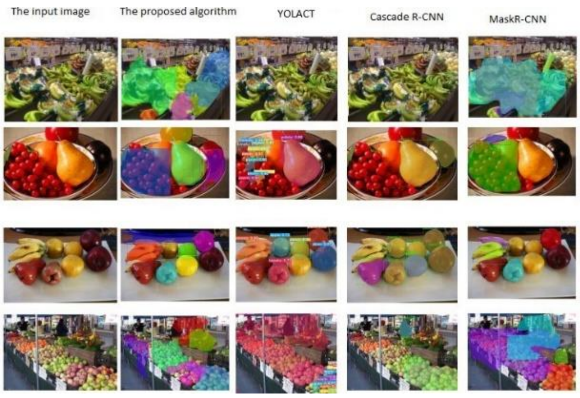

Figure 2. The visual testing findings of the proposed approach demonstrate the improved detection of numerous fruit instances when compared to existing algorithms

6. CONCLUSION

The present study attempts to develop deep learning methods for multiple fruit segmentation for high performance. This study uses deep learning to improve multiple fruit segmentation. Improvements to the ResNet-101 backbone improved multiple fruit segmentation. Selection of the right number of repetition blocks and filter size reduction in the early step achieved this. This method extracted more characteristics from input photos. Thus, this study generated an improved ResNet-101 backbone for multiple fruit segmentation. AP accuracy across thresholds is better with the proposed algorithm. This study's thresholds were higher than the three segmentation methods. With accuracy and training and testing time, the algorithm under consideration can quickly and accurately identify, detect, and segment various fruit items. However, this study used 27,000 photos from training and testing sets. Thus, to improve algorithm precision and resilience, more database images should be enlarged.

ACKNOWLEDGEMENT

This research is funded by the Deanship of Research and Graduate Studies in Zarqa University/Jordan.

REFERENCES

[1] N. Alshdaifat, M. A. Osman, and A. Z. Talib, "An improved multi-object instance segmentation based on deep learning," *Kuwait Journal of Science*, vol. 49, no. 2, 2022, doi: 10.48129/kjs.10879.

[2] A. Al-Momani, M. N. Al-Refai, S. Abuowaida, M. Arabiat, N. Alshdaifat, and M. N. A. Rahman, "The effect of technological context on smart home adoption in jordan," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 33, no. 2, p. 1186-1195, 2024, doi: 10.11591/ijeecs.v33.i2.pp1186-1195.

[3] B. Al-Naami, H. Abu Owida, M. Abu Mallouh, F. Al-Naimat, M. Agha, and A.-R. Al- Hinnawi, "A new prototype of smart wearable monitoring system solution for alzheimer's patients," *Medical Devices: Evidence and Research*, pp. 423–433, 2021, doi: 10.2147%2FMDER.S339855.

[4] Y. Li, H. Qi, J. Dai, X. Ji, and Y. Wei, "Fully convolutional instance-aware semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2359–2367, doi:10.48550/arXiv.1611.07709.

[5] K. He, G. Gkioxari, P. Dolla´r, and R. Girshick, "Mask r-cnn," in *Proc. IEEE Int. Conf. Comput. Vis*, 2017, pp. 2980–2988, doi: 10.1109/ICCV.2017.322.

[6] D. Bolya, C. Zhou, F. Xiao, and Y. J. Lee, "Yolact: Real-time instance segmentation," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 9157–9166, doi: 10.48550/arXiv.1904.02689.

[7] H. A. Owida, N. Alshdaifat, A. Almaghthawi, S. Abuowaida, A. Aburomman, A. Al-Momani,M. Arabiat, and H. Y. Chan, "Improved deep learning architecture for skin cancer classification," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 36, no. 1, pp. 501–508, 2024, doi: 10.11591/ijeecs.v36.i1.pp501-508.

[8] H. Owida, O. S. M. Hemied, R. S. Alkhawaldeh, N. F. F. Alshdaifat, And S. F. A. Abuowaida, "Improved deep learning approaches for covid-19 recognition in ct images," *Journal of Theoretical and Applied Information Technology*, vol. 100, no. 13, pp. 4925–4931, 2022.

[9] H. A. Owida, M. R. Hassan, A. M. Ali, F. Alnaimat, A. Al Sharah, S. Abuowaida, and N. Alshdaifat, "The performance of artificial intelligence in prostate magnetic resonance imaging screening," *International Journal of Electrical and Computer Engineering*, vol. 14, no. 2, p. 2234 – 2241, 2024, doi: 10.11591/ijece.v14i2.pp2234-2241.

[10] H. A. Owida, G. AlMahadin, J. I. Al-Nabulsi, N. Turab, S. Abuowaida, and N. Alshdaifat, "Automated classification of brain tumor-based magnetic resonance imaging using deep learning approach." *International Journal of Electrical and Computer Engineering*, vol. 14, no. 3, 2024, doi:10.11591/ijece.v14i3.pp3150-3158.

[11] A. Y. Alhusenat, H. A. Owida, H. A. Rababah, J. I. Al-Nabulsi, and S. Abuowaida, "A secured multi-stages authentication protocol for iot devices," *Mathematical Modelling of Engineering Problems*, vol. 10, no. 4, pp. 1352–1358, 2023, doi: 10.18280/mmep.100429.

[12] L. Deng, D. Yu et al., "Deep learning: methods and applications," *Foundations and trends® in signal processing*, vol. 7, no. 3–4, pp. 197–387, 2014, doi: 10.1561/2000000039.

[13] W. Shimoda and K. Yanai, "Cnn-based food image segmentation without pixel-wise annotation," i*n New Trends in Image Analysis and Processing–ICIAP 2015 Workshops: ICIAP 2015 International Workshops, BioFor, CTMR, RHEUMA, ISCA, MADiMa, SBMI, and QoEM, Genoa, Italy*, September 7-8, 2015, Proceedings 18. Springer, 2015, pp. 449–457, doi: 10.1007/978-3-319-23222-5 55.

[14] K. J. Pfisterer *et al.*, "Automated food intake tracking requires depth-refined semantic segmentation to rectify visual-volume discordance in long-term care homes," *Scientific Reports*, vol. 12, no. 83, 2022, doi: 10.1038/s41598-021-03972-8.

[15] D. Yang, Y. Du, H. Yao, and L. Bao, "Image semantic segmentation with hierarchical feature fusion based on deep neural network," *Connection Science*, vol. 34, no. 1, pp. 1772–1784, 2022, doi: 10.1080/09540091.2022.2082384.

[16] [16] V. Hemamalini, S. Rajarajeswari, S. Nachiyappan, M. Sambath, T. Devi, B. K. Singh, and A. Raghuvanshi, "Food quality inspection and grading using efficient image segmentation and machine learning-based system," *Journal of Food Quality*, vol. 2022, no. 1, p. 5262294, 2022, doi: 10.1155/2022/5262294.

[17] E. Aguilar, B. Remeseiro, M. Bolan˜os, and P. Radeva, "Grab, pay, and eat: Semantic food detection for smart restaurants," I*EEE Transactions on Multimedia*, vol. 20, no. 12, pp. 3266–3275, 2018, doi: 10.1109/TMM.2018.2831627.

[18] T. Le, "Mask r-cnn with data augmentation for food detection and recognition," *Computer Science, Agricultural and Food Sciences*, 2020, doi: 10.36227/techrxiv.11974362.

[19] Y. Li, X. Xu, and C. Yuan, "Enhanced mask r-cnn for chinese food image detection," *Mathematical Problems in Engineering*, vol. 2020, no. 1, p. 6253827, 2020, doi: 10.1155/2020/6253827.

[20] R. D. Yogaswara, E. M. Yuniarno, and A. D. Wibawa, "Instance-aware semantic segmentation for food calorie estimation using mask r-cnn," *in 2019 International Seminar on Intelligent Technology and Its Applications (ISITIA). IEEE*, 2019, pp. 416–421, doi: 10.1109/ISITIA.2019.8937129.

[21] S. Naritomi and K. Yanai, "Caloriecaptorglass: Food calorie estimation based on actual size using hololens and deep learning," *in 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW). IEEE*, 2020, pp. 818–819, doi: 10.1109/VRW50115.2020.00260.

[22] C.-H. Hsia, T.-H. W. Chang, C.-Y. Chiang, and H.-T. Chan, "Mask r-cnn with new data augmentation features for smart detection of retail products," *Applied Sciences*, vol. 12, no. 6, p. 2902, 2022, doi: 10.3390/app12062902.

[23] W. Jia, J. Wei, Q. Zhang, N. Pan, Y. Niu, X. Yin, Y. Ding, and X. Ge, "Accurate segmentation of green fruit based on optimized mask rcnn application in complex orchard," *Frontiers in Plant Science*, vol. 13, p. 955256, 2022, doi: 10.3389/fpls.2022.955256.

[24] S. P. Mohanty, G. Singhal, E. A. Scuccimarra, D. Kebaili, H. He´ritier, V. Boulanger, and M. Salathe´, "The food recognition benchmark: Using deep learning to recognize food in images," *Frontiers in Nutrition*, vol. 9, p. 875143, 2022, doi: 10.3389/fnut.2022.875143.

[25] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dolla´r, and C. L. Zitnick, "Microsoft coco: Common objects in context," *in Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13. Springer*, 2014, pp. 740–755, doi: 10.1007/978-3-319-10602-1 48.

[26] S. F. A. Abuowaida and H. Y. Chan, "Improved deep learning framework for multi food instance segmentations," *International Journal*, vol. 9, no. 4, 2020, doi: 10.30534/ijatcse/2020/308942020.

[27] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *in Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778, doi: 10.48550/arXiv.1512.03385.

[28] Y. Camgo¨zlu¨ and Y. Kutlu, "Analysis of filter size effect in deep learning," *Journal of Artificial Intelligence with Applications*, vol. 1, no. 1, pp. 20-29, 2020, doi: 10.48550/arXiv.2101.01115.

[29] W. S. Ahmed and A. a. A. Karim, "The impact of filter size and number of filters on classification accuracy in cnn," *in 2020 International Conference on Computer Science and Software Engineering (CSASE)*, 2020, pp. 88–93, doi: 10.1109/CSASE48920.2020.9142089.

BIOGRAPHIES OF AUTHORS

Esraa Abu Elsoud received the B.Sc. degree in Electrical Engineering from The Hashemite University, Jordan, in 2013 and M.Sc. in Cyber Security from The Hashemite University in 2023. Eng. Esraa's current research interests include cyber security, machine learning, big data and mobile network. She is currently a Lecturer in Zarqa University, Zarqa, Jordan. She can be contacted at email: eabuelsoud@zu.edu.jo.

Omar Alidmat received his B.Sc. degree in Computer Information Systems, M.Sc. degree in Computer Science from Al al-Bayt University, Jordan, and the Ph.D. degree in Computer Science from Universiti Sains Malaysia, Malaysia. Currently, Dr. Alidmat serves as an Assistant Professor in the Computer Science Department at Zarqa University, Jordan. His Ph.D. research focused on Artificial Intelligence, specializing in modelling crowd movements in panic situations, such as fires. He can be contacted at email: omar.alidmat@gmail.com**.**

Suhaila Abuowaida received the B.Sc. degrees in computer information system, M.Sc. degrees in computer science from AL alBayt University, Jordan, in 2012 and 2015, respectively, and the Ph.D. degree in Computer Science from Universiti Sains Malaysia, Malaysia, in 2023. Her research interests in-clude deep learning, depth estimation, point cloud and computer vision. She can be contacted at email: suhilaowida@gmail.com.

Esraa Alhenawi is a dedicated academic with a distinguished background in computer science. She earned her Ph.D. in computer science with honors in 2022, following her master's degree in the same field in 2015. She also holds a bachelor's degree in computer engineering from 2010, currently serving as an assistant professor at Zarqa University since September 2023. She can be contacted at email: ealhenawi@zu.edu.jo.

Nawaf Alshdaifat received the B. Sc degrees in computer science from AL al-Bayt University and M.Sc. degrees in computer science from the University of Jordan, Jordan, in 2002 and 2011, respectively, and the Ph.D. degree in Computer Science from Universiti Sains Malaysia, Malaysia, in 2023. His research interests include deep learning and machine learning. He can be contacted at email: n_alshdaifat@asu.edu.jo.

Ahmad Aburomman is currently a Full Professor of Artificial Intelligence in the Department of Data Science and Artificial Intelligence in the Zarqa University. He received the B.Sc. degree in Information Technology from Al-Balqa Applied University (BAU), Jordan, in 2005, and the M.Sc. degree in information technology (e-Services) from Griffith University, Australia, in 2009. He got the Ph.D. degree in time prediction in running business processes with the Research Center in Intelligent Technologies, University of Santiago de Compostela (CiTIUS). He can be contacted at email: aaburomman@zu.edu.jo.

Huah Yong Chan is a senior lecturer in the School of Computer Sciences, Universiti Sains Malaysia (USM). He is actively involved in grid computing research activities, both at the national and international level. He is also one of the project leaders of the e-Science Grid pioneer project in Malaysia, funded by the Ministry of Science, Technology, and Innovation. His research spans from grid computing to more specific issues such as resource allocation, load balancing, software agents, and middleware engineering. He can be contacted at email: hychan@usm.my.