Proceedings of the International Conference on Sustainable Computing and Data Communication Systems (ICSCDS-2022)
IEEE Xplore Part Number: CFP22AZ5-ART; ISBN: 978-1-6654-7884-7

A Hybrid Network Based on GAN and CNN for Food Segmentation and Calorie Estimation

R. Jaswanthi
Department of CSE
VR Siddhartha Engineering College
Vijayawada, India
jaswanthiregalla2655@gmail.com

E. Amruthatulasi
Department of CSE
VR Siddhartha Engineering College
Vijayawada, India
eliamruthatulasi@gmail.com

Ch. Bhavyasree
Department of CSE
VR Siddhartha Engineering College
Vijayawada, India
bhavyasreechirakala@gmail.com

Ashutosh Satapathy
Department of CSE
VR Siddhartha Engineering College
Vijayawada, India
ashutosh@vrsiddhartha.ac.in

2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS) | 978-1-6654-7884-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICSCDS53736.2022.9760831

Abstract— **Calories play an essential role in health aspects that lead to diseases like coronary heart disease, liver disease, cancer, and cholesterol. A study from 2020 reported that globally, overweight adults outnumber underweight individuals by more than 1.9 billion, while obese adults outnumber underweight ones by 650 million. Statistics from India show that abdominal obesity is the most significant risk factor, and it varies from 16.9% to 36.3%. Deep learning is an advanced image processing technology that solves problems and ensures food challenges because deeper networks have a better ability to process many features in an image. In our study, we propose a hybrid framework to predict the calorie content of food items on a plate. This includes three main parts: segmentation to segment the food from the image, image classification for classifying the food items, and calculating the calories present in those food items. A generative adversarial network is used for the segmentation, while a convolutional neural network is used for the classification and calorie estimation. The above models trained on the food images from the UNIMIB 2016 dataset have correctly recognized and estimated the calories of a food item with an accuracy of 95.21%.**

Keywords— Image segmentation, Image classification, Calorie estimation, Generative Adversarial Network, Convolutional Neural Network.

I. INTRODUCTION

In today's generation, there is a significant increase in the obesity rates for both adults and children. Foods that we eat that are high in calories will lead to obesity. It turns out to be a risk factor. According to the WHO, 3 out of 5 men and 6 out of 10 women are obese. As consulting a dietician is difficult in every situation, there will be a need for AI-based food segmentation and recognition techniques to identify food items. Focusing on a specific situation where only one food item is placed on a unique colour palette, the plate is separated from the background, and the food is separated from the container. Here, a Generative Adversarial Network (GAN) trained on a mid-sized food dataset appropriately extracts regions of interest (RoI) from an RGB image. After the segmentation, a Convolutional Neural Network (CNN) trained on the same dataset correctly classifies the RoI as a food as a food item and estimates the calories within it.

Image segmentation is commonly used to find objects and boundaries in an image by splitting an image into several parts, which are called segments. An image segmentation algorithm splits and groups a particular set of pixels of an image, assigns distinct labels to pixels, and pixels with the samelabels uniquely depict an object. These labels could be used to define borders, draw lines, and distinguish the essential food items from the background image [1]. The use of deep learningtechniques in the food science domain is highly successful dueto its high potential for autonomous feature learning, primarily for food category classification and food calorie estimation. The segmented food items from their backgrounds using a novel segmentation algorithm trained on CNN-based features are shown in Fig. 1 [2].

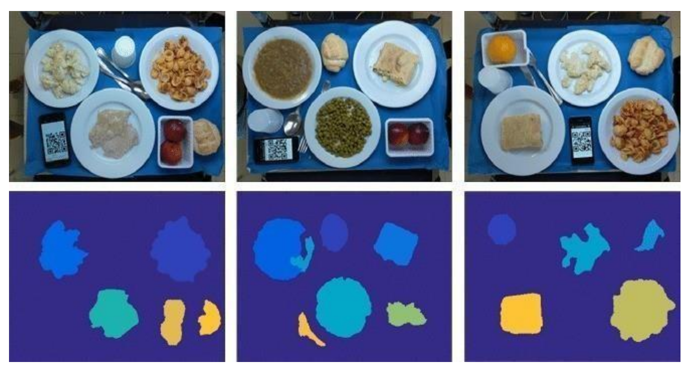

Fig.1. Segmented image samples of food images from the UNIMIB2015 dataset.

Object recognition is one of the important techniques in the field of pattern recognition that allows a machine to understand and categorise an object from a group of similar or different objects. Multiple object recognition techniques are employed to recognize the food item and estimate the number of calories present in it. A generic-based food recognition technique takes the entire food region into consideration while extracting the border information such as edges, lines, and curves. However, a structural matching method considers the peel color and texture, the shape and size of a food item, and its location to obtain various geometrical features for food item classification and calorie estimation. A hybrid recognition approach performs better than generic and structural-based recognition techniques as it utilizes both generic and structural features to classify a different food items [3]. Like a structural matching method, geometric-based recognition technique successfully extracts meaningful features from peel color and texture, as well as the shape and size of a food item with different orientations in the presence of occlusion at different lighting conditions. Supervised pattern recognition models such as Support Vector Machine (SVM), evolutionary algorithms, and neural networks trained on the geometric feature vectors of different food images have achieved remarkable accuracy. Recognition of different food items present in the image samples from the Foodseg103 dataset is exhibited in Fig. 2 [4].

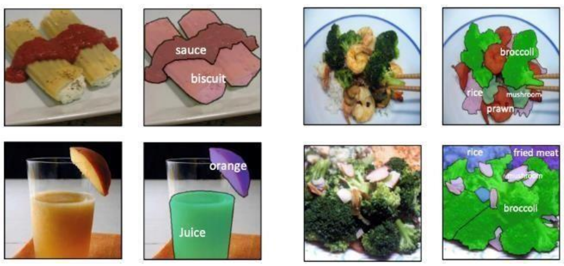

Fig.2. Recognition of the food items from the Foodseg103 dataset image samples.

Generative modeling is an unsupervised learning task in machine learning that involves automatically discovering and learning the regularities in input data in such a way that the model can be used to generate new examples. A GAN contains two sub-models: the generator model trains to generate new samples, and the discriminator model tries to recognize samples as either real (from the domain) or fake (generated). As a GAN efficiently learns features from a set of training data from a small dataset by creating new data with the same characteristics from random noise as the training data, a GAN is a suitable choice for geometric based segmentation to accurately extract food items' regions from their backgrounds [5].

A CNN is an elegant and sophisticated deep learning technique for object detection, localization, and classification [6]. It has been employed in a wide range of real-world scenarios, including food science, activity recognition, motion detection, diagnosis, pose detection, and recognition. A CNN consists of multiple convolutional and pooling layers backed by a multi- layer fully connected neural network with a reasonable amount of padding and strides at convolutional layers and dropouts at fully connected layers [7]. Eventually, they made a CNN suitable for image processing and computer vision applications with greater accuracy. Any gradient descent optimizer, such as Adam, Stochastic Gradient Descent, or RMSProp, significantly optimizes the error rate with proper tuning of the weights of thisclassifier in each epoch. Here, CNN is utilized for classifying food images based on their geometrical features, as it exploits these spatial hierarchical features of data while classifying them into different classes.

In this paper, we have proposed a hybrid framework for food calorie estimation that is built on Pix2Pix GAN and ResNet50 CNN. So, we have introduced the Pix2Pix GAN and ResNet50 and the use of these models to segment the food regions from their background, classify individual food items, and estimate the calories present in those food items.

II. RELATED WORK

Ciocca, Napoletano, and Schettini created a diet monitoring system where the Hough transform detected plates in circle shapes in the image samples. Individual plate regions are segregated into multiple patches of size 40x40 pixels using a patch-based food recognition algorithm. The extracted features from each patch were supplied to a previously trained k-nearest neighbor (k-NN) classifier for classifying the patch as a food or non-food region. Subsequently, patches with bogus labels were adjusted to obtain uniform regions of the food items. A setof food items for a customer was fetched from the database, andthe ratio for each food item was computed by dividing the number of patches by the total number of patches before and after the meal, respectively [8]. Finally, the calories of leftoverfood items were estimated with the help of a predefined nutrienttable. The result shows that the proposed system correctlyestimated calories for 85% of the food images.

The Segmentator module applied binarization to extract RoIs from the saturation channel of an HSV gamma-corrected image. At the same time, the JSEG segmentation algorithm with region removal technique not only extracted RoIs from a color image but also removed spurious regions from the output image. The segmented images generated by both the processes were combined and cleaned using geometric constraints to produce final RoIs. The predictor module divided a RoI into multiple patches of size 140x140 pixels, extracted features from the patches and the whole RoI, and performed local and global classifications to decide the class label of the RoI. The tray analysis pipeline built on segmentator and predictor has obtained a food recognition accuracy of 78.3% on the UNIMIB 2016 dataset [9].

A deep CNN, so-called NutriNet, was developed by Mezgec and Seljak to recognize images containing food and drink items [10]. Unlike the previous models, NutriNet tries to identify RoIs in an RGB image based on the bounding box data provided with the image samples while training the model. Here, the NutriNet model was trained using the AdaGrad optimizer, and the result shows that the classification accuracy on the UNIMIB 2016 dataset was 86.39%. Minija and Emmanuel used Local Variation Segmentation (LVS) for segmenting the food items and multi-Kernel SVM for recognizing individual food items that helped them predict the calorie content of those food items [11]. A multi-Kernel SVM trained on 80% of the image samples in the UNIMIB 2016 dataset achieved a recognition rate of 94.7% on 20% of the same dataset's image samples.

The proposed Cauchy, Generalized T-Student, and Wavelet kernel-based Wu-and-Li Index Fuzzy Clustering (CSW-WLIFC) segmentation algorithm utilizes a kernel built on the CSW generated segmented images. The Whale Levenberg Marquardt Neural Network (WLM-NN) model was trained on features such as appearance, texture, and color from segmented regions for recognizing food items and estimating their calories [12]. The proposed model trained on the UNIMIB 2016 had obtained 96.27% classification accuracy after segmenting the food items using CSW-WLIFC, whose accuracy was 99.9%.

Minija and Emmanuel came up with a Deep Belief Network(DBN) based on the Imperialist Competitive Algorithm

(IpCA) for recognizing food items and estimating the presence of calories [13]. Shape, Wavelet, scattering transform, and histogram features were extracted from the segmented food regions generated by the Bayesian Fuzzy Clustering and give to the DBN for classification of food items and estimating the number of calories present in them. The experimental result shows that the proposed method trained to accurately recognize 98.77% of the food items from the UNIMIB 2016 image samples.

The initial implementation of NutriNet by Mezgec and Seljak performed pixel-level recognition and recognized one food itemper image. In this article, it was trained with a larger dataset to recognize multiple items in images that included foods and beverages [14]. At first, images were downloaded from the Internet related to 520 food and beverage items. A deep learning model removed spurious images were removed from the data whose size was increased further by applying augmentation such as scaling, rotation, translation, or adding noise to the images. The pixel level classification accuracy obtained by the NutriNet was 92.18%, compared to the classification accuracy obtained by the CNN.

Poply and Jothi used mask region-based CNN (R-CNN) for food item segmentation and recognition. The generated mask by mask R-CNN for a particular food item was used to compute the surface area of the food region in square inches, which was multiplied with the calories per square inch of the recognized food item to estimate the calories present in it. The UNIMIB2016 dataset was used to train and test the models, whichcontains 1027 RGB images across 73 food classes. The modelachieved an average accuracy of 95.45% and 93.4% while estimating the calories of a food item and whole meals consisted of multiple food items, respectively [15].

Segmentation of the food items was carried out using object detection and semantic segmentation based on CNN, where pixels of each food item were recognized after detection of food regions in an image. The identified pixels of a food item were used to estimate the volume and mass of the recognized food item [16]. The predicted mass, volume, and density information of the identified food item helped to compute the calories with the help of a calorie look-up table. The researchers applied transfer learning while training the models on the UNIMIB 2016, whose calorie estimation accuracy for a single food item and complete meals were 90.80% and 93.06%, respectively.

Pfisterer et al. developed the Automated Food Imaging and Nutrient Intake Tracking (AFINI-T) technique to track the amount of food intake by long-term care (LTC) inhabitants. It utilized RGB-D evaluation-based food segmentation and an auto-encoder for food item classification after training the model using the UNIMIB 2016 augmented image samples, whose top-1 classification accuracy was 88.9% [17]. The volume estimation error and error intake of the AFINI-T technology were 2.59.2 mL and -0.436.7 mL, respectively. The nutrients estimated from the volume by this technology were strongly correlated (0.92 - 0.99) with the estimated nutrients from the mass of the food items.

III. PROPOSED METHODOLOGY

Fig. 3 illustrates the flow diagrams of the hybrid network for image segmentation followed by food item recognition and calorie estimation. As explained earlier, GAN segregates food item regions from the image background, and the extracted food regions are supplied to CNN for food item recognition. The Shape from Silhouettes method estimates the volumes in cubic centimeters of the segmented food regions [18], which are multiplied with their density values to determine their weights in grams and subsequent estimation of the calories present in those food items. For our implementation, the density and calorie information of different food items are taken from the foodinfo.us and CalorieKing.com websites respectively [19, 20].

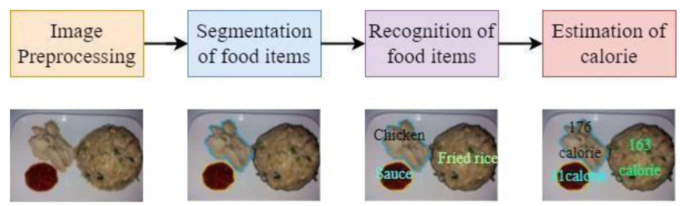

Fig.3. Flow diagram of hybrid framework for food calorie estimation

The process begins with the admin or developer's selection of appropriate image datasets that consist of a lot of spatial imagesof food items in steps 1 and 2 as shown in Fig. 4. Noise and redundant information present in those images are removed by applying smoothing and feature selection algorithms in step 3. By performing steps 4 to 9 repeatedly, we have found that the combination of the Pix2Pix GAN and ResNet50 generates high accuracy while estimating the calorie content of food items. After successful evaluation of the above models, the hybrid framework has been deployed in step 11, which is ready to takespatial food images and generate accurate results in steps 12 and 13. As there is a need to train the models with a new set of food image samples to improve the robustness of the proposed framework, new food datasets are collected; redundant information and noise are removed from these as shown in steps14, 15 and 16. The previously trained Pix2Pix GAN and ResNet50 learn from the new datasets using the incremental learning technique in step 17. In step 18, the framework is redeployed for estimating food calories from any random food image samples supplied by users in steps 19 and 20.

With the help of random noise, the generator creates new image samples like real images, and the discriminator tries to identify the likelihood between the real and generated images using a loss function. The loss value propels the generator and discriminator to adjust their parameters for generating identical images to a great extent and determining the real and fake images with minute differences, respectively. In Fig. 5, F and F' are the actual and generated image setsused to train a GAN. G(η) is the generating function thatproduces F' using random noise η, whereas ϕ(t) is the discriminating function that takes F or F' as input and classifiesthem as real or fake images. Finally, m(ɵ) and n(ɵ) are the functions responsible for fine tuning the parameters of the generator and discriminator during the back propagation.

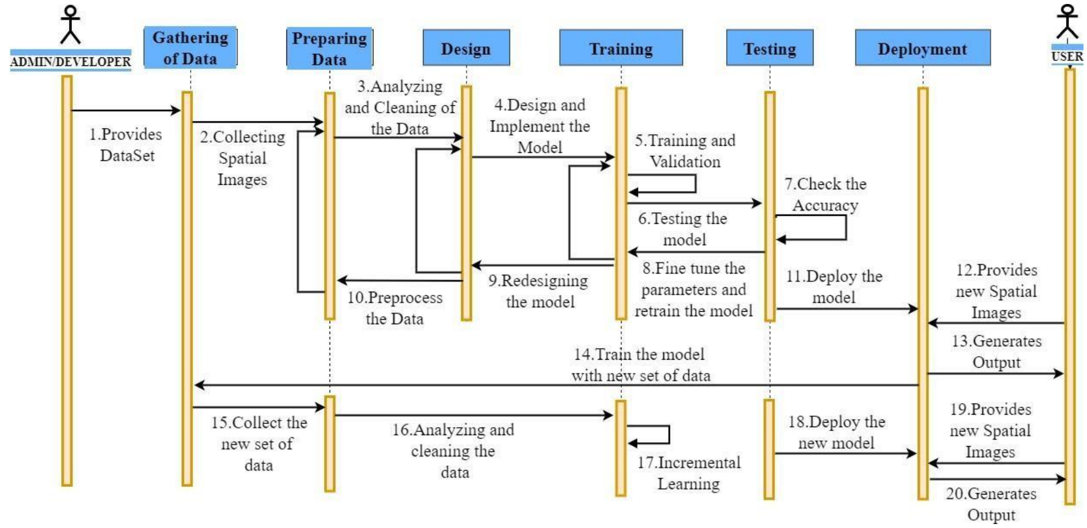

Fig. 4. Sequence diagram of the hybrid network based on GAN and CNN

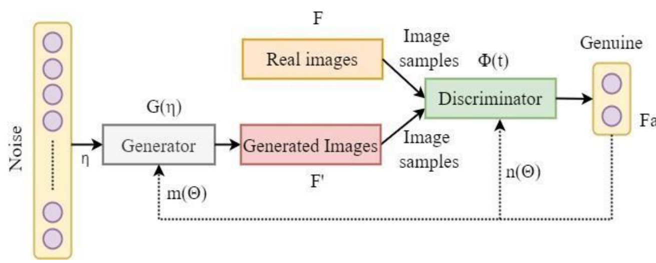

Fig.5. Fundamental architecture of a GAN.

Pix2Pix GAN is meant for image-to-image translation and works as a conditional GAN (CGAN) that generates images like training images based on the label supplied to it. The generated image seems reasonable for the desired domain, and it is a reasonable translation of the supplied input. The generator and discriminator are the CNNs made of layers mostly found in all CNNs, such as convolution, batch normalization, ReLU, concatenate, Eltwise, and pooling layers. Here, the generator uses UNet instead of a traditional autoencoder for image translation, and the discriminator uses PatchGAN for estimating similarity between the ground truth and generated images [21]. In our work, the actual images along with segmented images from the generator, the ground truth segmented images, are supplied to the discriminator to distinguish between real and fake, as shown in Fig. 6.

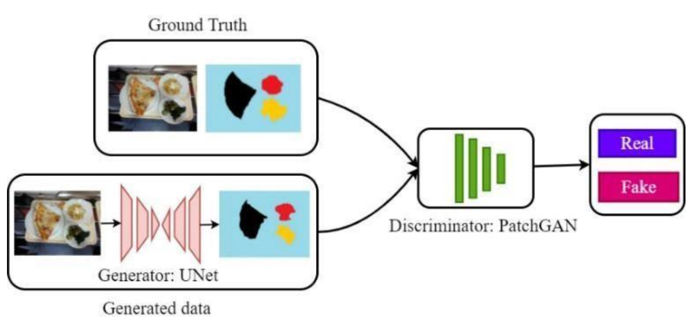

Fig.6. Pix2Pix GAN for food item segmentation

ResNet50 is made up of forty-nine (one 7x7, thirty-two 1x1, and sixteen 3x3) convolutional layers, one max pool, one average pool, and ends with a fully connected layer along witha softmax function [22]. As kernels at the pooling layer are notembedded with any weight values, the parameter size is decreased without any fine tuning of parameters. Apart from that, each convolution layer is followed by a batch- normalization and ReLU layers, which are responsible for normalizing the features and mapping them to +ve values, which capture the most prominent features and substantially minimize the cost during learning. Due to unreachable or weak back propagation signals that make no changes in weight values at the initial convolutional layers, increasing CNN depth beyond a specific threshold raises the error rate, and optimization of many weights may lead to a high training error rate [23]. It is minimized by putting three consecutive convolutional layers (1x1, 3x3, 1x1) into each residue module (total: 16 modules) as shown in Fig. 7. All the residue modules are joined serially, where each residue module's output is the linear combination of its input and output from its final layer. In the end, the SoftMax loss function evaluates the loss after every epoch, which is used as a parameter by an optimization function to adjust the weights at each layer during backward computation.

IV. RESULTS AND DISCUSSION

For food recognition, segmentation, and calorie estimation, we used the UNIMIB 2016 database. It has 73 different types of cuisine and 1027 tray image samples with dishes from each of the 73 groups. The canteen on the campus of the University of Milano-Bicocca provided the data for this dataset. The setup isunique in that each image portrays various items on a tray, and some foods (such as fruit, bread, and desserts) are served on place mats rather than plates. Visual distortions as well as lighting shifts owing to shadows may be seen in the photographs. The Computer Vision Annotation Tool (CVAT) [24] was used to establish the ground truth for these

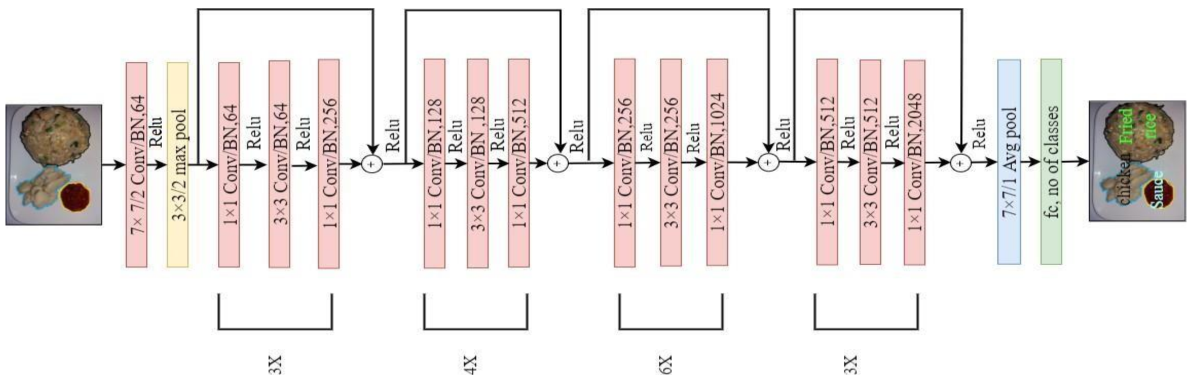

Fig.7. ResNet-50 architecture for recognizing food items from segmented images.

images, which were needed to train Pix2Pix GAN.

As discussed earlier, after successful recognition of food items,ROIs of those foods were used to estimate the volume in cc using the Shape from Silhouettes method. The mass of a food item was calculated by multiplying the density (g/cc) of the food with its estimated volume. The total calorie amount is the multiplication of calories per 100 grammes and the mass of the identified food as given in equation 1.

$$\text{total\_calories} = (\text{calories}/100) * \text{food\_mass} \quad (1)$$

The experiment was conducted on a Google Colab notebook running on the top of a single-core hyper-threaded Xeon @2.3GHz CPU, an Nvidia Tesla K80 GPU, and 12.7 GB of RAM, with a virtual machine having a disc space of approximately 100 GB. Keras is a Python-based deep learning API that runs on top of the TensorFlow machine learning As Fig. 8 shows, the Pix2Pix GAN has properly segmented the RoI from the background, and ResNet50 has successfully recognized the food as banana.

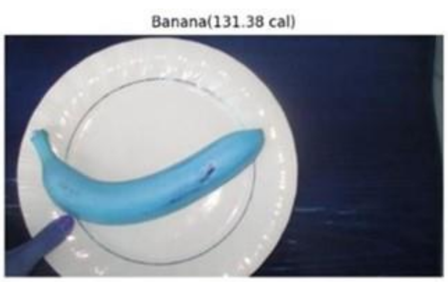

Fig.8. Estimated calories from banana

As Fig. 8 shows, the Pix2Pix GAN has properly segmented the RoI from the background, and ResNet50 has successfully recognized the food as banana. Fig. 9 shows, the Shape from Silhouettes algorithm had estimated 157.04 cc as the volume of the banana from the highlighted RoI in Fig. 8. Food mass obtained by multiplying 157.04 cc with the density value (g/cc) of a rip banana provided in the foodinfo.us website. Finally, the calorie amount in Fig. 8 food-item was calculated by multiplying 147.62 g mass by thecalories per gram.

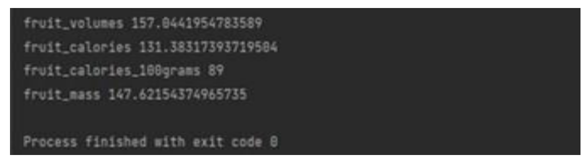

Fig.9. Food volume and calorie calculation of banana in Fig. 8.

$$e = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}|p'_{ij}-p_{ij}|}{\sum_{i=1}^{m}\sum_{j=1}^{n}p_{ij}} * 100 \quad (2)$$

In equation 2, e is the error percentage computed over n food items across m images. Here, n is the number of food items present in the ith image. p'_{ij} and p_{ij} are the estimated and actual calories of the jth food item in the ith image. The overall accuracy of the model is calculated by subtracting e from 100.

TABLE I. Comparison between the previous implementations and proposed framework based on food recognition or calorie estimation accuracies.

Related Work	Recog. Task	Accuracy (%)
Ciocca, Napoletanoand Schettini [7]	Calorie	85
Ciocca, Napoletanoand Schettini [8]	Food	78.3
Mezgec and Koroušić [9]	Food	86.39
Minija and Emmanuel [10]	Food	94.7
Minija and Emmanuel [11]	Food	96.27
Minija and Emmanuel [12]	Food	98.77
Mezgec and Koroušić [13]	Food	92.18
Poply and Jothi [14]	Calorie	95.45
Poply and Jothi [15]	Calorie	90.80
Pfisterer et al. [16]	Food	88.90
Our hybrid network	Calorie	95.21

Table I displays the food recognition and calorie estimation accuracy of previous implementations. Our network provides better results than most of the implementations, except for the models developed by Minija and Emmanuel and Poply and Jyothi. The food recognition accuracy of models by Minija and Emmanuel was 96.27% and 98.77%, whereas the 95.45% calorie estimation accuracy of the model by Poply and Jyothi was 96.27%. As calorie estimation is carried out after food recognition, we can say our framework performs better than the previous implementations, except for Poly and Jyothi's calorie estimation model.

V. CONCLUSION

In this article, we have proposed a hybrid network built on GANand CNN for food item segmentation, recognition, and calorie estimation. Pix2Pix GAN extracted food RoIs, while ResNet50classified the extracted food RoI to a certain class. Lastly Shapefrom Silhouettes method estimated the volume from a food region for mass calculation by multiplying density value to it. The final estimated calorie was the multiplication of the mass and precomputed calories per gram obtained from online resources. The results shows that the proposed hybrid network achieved 95.21% calorie estimation accuracy superior to most of the previous implementation. In future, the proposed networkmust be improved further to predict calories of multiple food items including fruits and vegetables on a plate.

REFERENCES

[1] S. Minaee et al. “Image segmentation using deep learning: A survey,” *IEEE Trans. Pattern Anal. Mach. Intell.,* 2021.

[2] G. Ciocca, P. Napoletano and R. Schettini, “Food recognition: a new dataset, experiments, and results,” *IEEE J. Biomed. Health Informat.,* vol. 21, no. 3, pp. 588-598, 2016.

[3] D. N. Parmar and B. B. Mehta, “Face recognition methods and applications,” *arXiv preprint arXiv:1403.0485,* 2014.

[4] X. Wu et al. “A large-scale benchmark for food image segmentation,” in *Proceedings of the 29th ACM Int. Conf. Multimedia*, 2021, pp. 506-515.

[5] A. Radford, L. Metz and S. Chintala, “Unsupervised representation learning with deep convolutional generative adversarial networks,” *arXiv preprint arXiv:1511.06434,* 2015.

[6] P. Kim, "Convolutional neural network," in *MATLAB deep learning*, Berkeley, CA, Apress, 2017, pp. 121-147.

[7] A. Bhaskar, “Survey on evolving deep learning neural network architectures,” *J. Artif. Intell.*, vol. 1, no. 2, pp. 73-82, 2019.

[8] G. Ciocca, P. Napoletano and R. Schettini, “Food recognition and leftover estimation for daily diet monitoring,” in *Int. Conf. Image Anal. Process.*, Springer, 2015, pp. 334-341.

[9] G. Ciocca, P. Napoletano and R. Schettini, “Food recognition: a new dataset, experiments, and results,” *IEEE J. Biomed. Health Informat.,* vol. 21, no. 3, pp. 588-598, 2016.

[10] S. Mezgec and B. K. Seljak, “NutriNet: a deep learning food and drink image recognition system for dietary assessment,” *Nutrients,* vol. 9, no.7, p. 657, 2017.

[11] S. J. Minija and W. S. Emmanuel, “Image processing-basedClassification and Segmentation using LVS based Multi-Kernel SVM,”in *2018 Int. Conf. Smart Syst Inventive Technol. (ICSSIT)*, IEEE, 2018, pp. 61-65.

[12] W. R. S. EMMANUEL and S. J. MINIJA, “Fuzzy clustering and Whale-based neural network to food recognition and calorie stimationfor daily dietary assessment,” *Sādhanā,* vol. 43, no. 5, pp. 1-19, 2018.

[13] S. J. Minija and W. R. S. Emmanuel, “Imperialist competitive algorithm-based deep belief network for food recognition and calorie estimation,” *Evol. Intell.,* pp. 1-16, 2019.

[14] S. Mezgec and B. K. Seljak, “Using deep learning for food and beverage image recognition,” in *2019 IEEE Int. Conf. Big Data (Big Data)*, IEEE, 2019, pp. 5149-5151.

[15] P. Poply and A. A. Jothi J, “An Instance Segmentation approach to Food Calorie Estimation using Mask R-CNN,” in *Proc. the 2020 3rd Int. Conf. on Signal Process. Mach. Learn*, 2020, pp. 73--78.

[16] P. Poply and J. A. A. Jothi, “Refined Image Segmentation for Calorie Estimation of Multiple-dish food items,” in *2021 Int. Conf. Comput. Commun. Intell. Syst. (ICCCIS)*, IEEE, 2021, pp. 682-687.

[17] K. J. Pfisterer et al. “Enhancing Food Intake Tracking in Long-Term Care withAutomated Food Imaging and Nutrient Intake Tracking (AFINI-T) Technology,” *arXiv preprint arXiv:2112.04608,* 2021.

[18] C. Xu et al. “Image- based food volume estimation,” in *Proc. the 5th Int. workshop on Multimedia for cooking and eating activities*, 2013, pp. 75-80.

[19] “Food densities (weight / volume relationships of selected food items),” *Food Densities (weight / volume relationship)*. [Online]. Available: http://foodinfo.us/Densities.aspx. [Accessed: 23-Jan-2022].

[20] *Food Nutrition Facts and Free Calorie Counter | CalorieKing*. [Online]. Available: https://www.calorieking.com/us/en/. [Accessed: 23-Dec-2021].

[21] P. Isola, J. Y. Zhu, T. Zhou and A. A. Efros, “Image-to-image translation with conditional ad versarial networks,” in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, 2017,pp. 1125-1134.

[22] K. He, X. Zhang, S. Ren and J. Sun, “Deep residual learning for image recognition,” in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, 2016, pp. 770-778.

[23] J. Kolbusz, P. Rozycki and B. M. Wilamowski, “The study of architecture MLP with linear neurons in order to eliminate the “vanishing gradient” problem,” in *Int. Conf. Artif. Intell. Soft Comput.*, Springer, 2017, pp.97-106.

[24] U. S. Hansen, E. Landau, M. Patel and B. Hayee, “Novel artificial intelligence-driven software significantly shortens the time required for annotation in computer vision projects,” *Endoscopy Int. Open,* vol. 9, no. 4, pp. E621-E626, 2021