Calorie Counter: Recognition of food products, ready meals and their caloric values using neural networks

Poltavskiy A.V.

Msc fellow, Southern Federal University, Russia

poltavsky@sfedu.ru

poltavskypro@gmail.com

***Abstract*— Automatic image-based food recognition is a particularly challenging task. Traditional approaches to image analysis have in the past provided poor classification accuracy, while deep learning approaches have made it possible to identify food types and their ingredients. The contents of food dishes are usually deformable objects, usually involving complex semantics, which makes the task of determining their structure extremely difficult. Deep learning techniques have already shown very promising results in such tasks, so this paper focuses on presenting some of the popular approaches and techniques used in image-based food recognition.**

***Keywords*— Food recognition, calorie count, image segmentation, salient object detection**

I. Introduction

Despite the presence of dozens of applications, algorithms and systems, the problem of recognizing dishes (food products) and their ingredients has not been completely solved by the expert communities of machine learning and computer vision. The main current limitation of health and nutrition tracking apps is the need to manually enter each meal.

Neural network approaches from researchers and leaders in the calorie counting segment make it easier to track food intake based on image classification. The systems recognize the class of the object, and as a result, they give the user its caloric content and characteristics. The difficulty in solving this problem is the number of classes of dishes and their similarity in some cases.

However, this recognition does not answer the main question: “How many calories are on my plate?”. This work is devoted to its deep study and solution.

.

II. Deep learning in solving classification and segmentation problems

Speaking of deep learning for solving classification and segmentation problems, training the latter requires an input mask of an object with the corresponding class, when the classifier is only a class. However, using the master object detection model, comparable accuracy in object mask detection and class recognition is demonstrated. In accordance with the model output described in Figure 1, the coupling of the main object detection with the image classification by mask makes it possible to train the segmentation model for a large cluster of noticeable objects and performs similar functions.

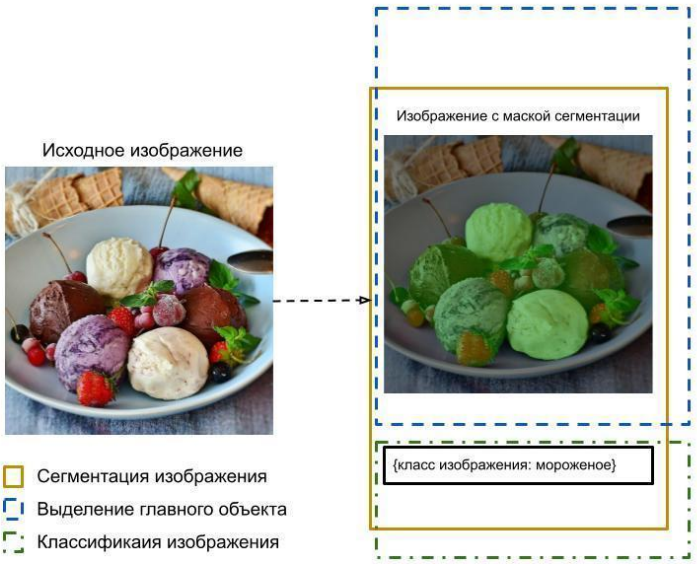

Fig. 1 Sample images of the food-101 dataset

This approach allows you to automate the process of data markup for training neural network segmentation models and can be used as such.

The complexity of training a neural network segmentation model is revealed by a specific example - the image domain. Due to the complexity of mask markup, public datasets such as food-101 and others [5] include only the image class, examples of image classes are shown in Figure 2.

Fig. 2 Sample images of the food-101 dataset

III. Limitations of health and nutrition tracking apps

The main current limitation of health and nutrition tracking apps is the need to manually enter each meal or scan a barcode. This approach has its advantages in the form of determinism by default, but forces the user to go through a monotonous process of selecting a dish from the database. And neural network approaches from researchers and leaders of the calorie counting segment simplify the tracking of food consumed based on the classification of images, followed by adjusting the ingredients of the found dish as shown in Figure 3.

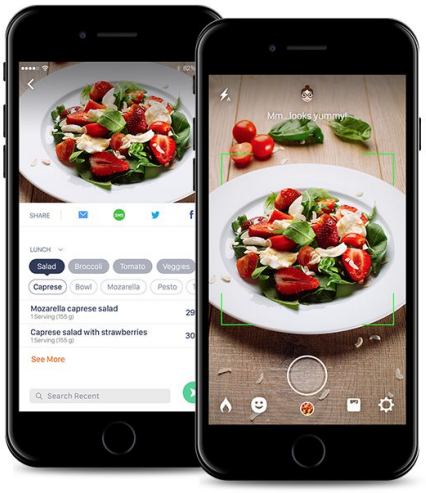

Fig. 3 calorie mama app layout

The systems recognize the class of the object, and as a result, they give the user its caloric content and characteristics. The difficulty in solving this problem is the number of classes of dishes and their similarity in some cases Figure 4.

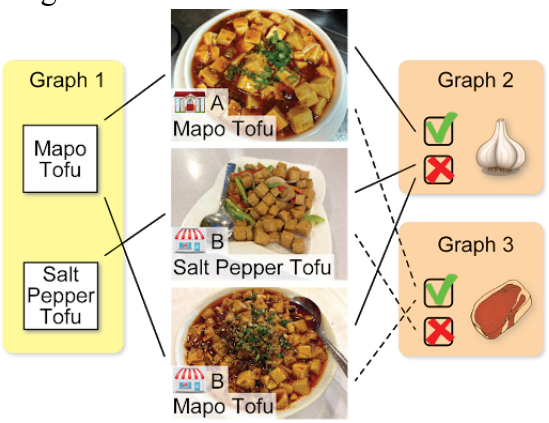

Fig. 4 Classification of different types of tofu dishes

IV. Count calories from a photo

The implementation of the task of determining calories from a photo can be presented in the form of the following points:

1) Using the database, we determine the dish from the photo
2) By dish, nutritional value: caloric content, bzhu, related metrics
3) By the segmentation mask relative to the plate, we determine the volume.

4) By volume, using the summary density tables, we find the mass of the object
5) By the weight of the dish, we calculate the exact calorie content

Email address is compulsory for the corresponding author.

A. Calculating the volume/mass of a product from a photo

The calculation of the volume is reduced to the problem of converting the density of the body and the known dimensions. The calculation of body mass is expressed in terms of density and volume by the following formula: body mass is the product of the density of the substance that makes up the body by its volume.

$$m=pV$$

Formula 1. V is the volume of the body (cubic meters), m is the mass of the body (kg), p is the density of the substance (kg/cubic meters)

B. Transformation by density and known dimensions

Used heuristics for density transformation and known / assumed dimensions:

1) Calculating the size of the plate. On average, it is unified and predictable. The user can easily choose from the existing widespread dimensions of Fig. 8 (15, 22, 30 cm.)
2) Calculating the relative area of the occupied dishes to the surface of the plate Fig. 9.
3) Comparing the two calculations and obtaining the actual area and calculating the third dimension (height) for the predicted dish (pizza: 2-3cm, cake: 7-10 cm).
4) Use the density of the product to calculate its maximum weight according to formula 1.

Fig. 5 Example of calculating the size of the object on which the dish is located and its relative area.

V. Conclusions

An important aspect of health monitoring is the effective management of

journal of food consumption. It can help in the treatment of diseases related to nutrition, such as obesity, diabetes, and even cardiovascular diseases. In addition, food accounting can help fitness enthusiasts and people who want to reach a target weight. However, food logging is cumbersome and requires not only additional effort to record regularly consumed food, but also sufficient knowledge about the types of food consumed, which is difficult due to the presence of a widely represented selection.

With the growing dependence on smart devices, we are using the convenience of input offered by smartphones, and in this paper, we presented an algorithm and a prototype for developing a system for automatically determining calories in food and ready meals.

Due to the complexity of marking up dish masks, an approach based on the composition of neural network models of classification and determination of the main object in the image, which serves as segmentation, was applied. The proposed method of recognition of dishes based on the detection of the main object with the classification of the image by the mask allows you to achieve more impressive results, thanks to the improved quality of classification and clarity of the recognition process. The classification quality is improved by a more perfect input image-a mask of the selected object, instead of an entire photo with a lot of objects or unnecessary, distracting details. The clarity of the results and the potential feedback from the user is that the user specifically sees the classification object and its mask - in the case of incorrect or inaccurate recognition, the user can correct the mask or class of the object, which will allow using the method of training the model throughout the life cycle (continuous learning).

A database of ingredients and ready meals was also established, which resulted in not only the number of calories per certain amount of product, but also other nutritional characteristics, such as macronutrients and BZHU analysis. Using heuristics, neural network models of assumption and user interaction, a method for determining the mass of a dish was proposed.

All three components ensure the successful construction of a system for automatic recognition of calories of dishes and food products on plates/trays, and so on. The system can expand the type of supported use cases and the accuracy of their recognition thanks to user feedback.

References

[1] Haralick, R.M. and Shapiro, L.G., 1985. Image segmentation techniques. Computer vision, graphics, and image processing, 29(1), pp.100-132.

[2] Lu, D. and Weng, Q., 2007. A survey of image classification methods and techniques for improving classification performance. International journal of Remote sensing, 28(5), pp.823-870.

[3] Borji, A., Cheng, M.M., Jiang, H. and Li, J., 2015. Salient object detection: A benchmark. IEEE transactions on image processing, 24(12), pp.5706-5722.

[4] Qin, X., Zhang, Z., Huang, C., Dehghan, M., Zaiane, O.R. and Jagersand, M., 2020. U2-Net: Going deeper with nested U-structure for salient object detection. Pattern Recognition, 106, p.107404.

[5] Bossard, L., Guillaumin, M. and Van Gool, L., 2014, September. Food-101–mining discriminative components with random forests. In European conference on computer vision (pp. 446-461). Springer, Cham.

[6] London, M. and Smither, J.W., 1999. Empowered self‐development and continuous learning. Human Resource Management: Published in Cooperation with the School of Business Administration, The University of Michigan and in alliance with the Society of Human Resources Management, 38(1), pp.3-15.

[7] J. Han, D. Zhang, G. Cheng, N. Liu, and D. Xu, “Advanced deep-learning techniques for salient and category-specific object detection: a survey,” IEEE Signal Processing Magazine, vol. 35, no. 1, pp. 84–100, 2018.

[8] A. Borji, “Saliency prediction in the deep learning era: Successes and limitations,” IEEE Trans. Pattern Anal. Mach. Intell., 2019

[9] A. Borji, M.-M. Cheng, Q. Hou, H. Jiang, and J. Li, “Salient object detection: A survey,” Computational Visual Media, pp. 1–34, 2019.

[10] J. Shi, Q. Yan, L. Xu, and J. Jia, “Hierarchical image saliency detection on extended cssd,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 38, no. 4, pp. 717–729, 2015.C. Yang, L. Zhang, H. Lu, X. Ruan, and M. H. Yang, “Saliency detection via graph-based manifold ranking,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2013, pp. 3166–3173.

[11] W. Wang, J. Shen, L. Shao, and F. Porikli, “Correspondence driven saliency transfer,” IEEE Trans. Image Process., vol. 25, no. 11, pp. 5025–5034,

[12] F. Guo, W. Wang, J. Shen, L. Shao, J. Yang, D. Tao, and Y. Y. Tang, “Video saliency detection using object proposals,” IEEE Trans. Cybernetics, 2017