Scientific and Social Research, 2021, Volume 3, Issue 4
http://journals.innosciencepress.com/index.php/ssr
ISSN Online: 2661-4332

An 11-Categorical AI Food Classification Model Based on Mobile-Net Neural Networks

Yongqing Yu[1], Yishan Zou[2], Yu Sun[3]*

[1]BASIS International School Park Lane Harbour, Huizhou 516001, China

[2]Department of Education, University of Pennsylvania, Philadelphia 19104, USA

[3]Department of Computer Science, California State Polytechnic University, Pomona 91768, USA

***Corresponding author:** Yu Sun, yusun@cpp.edu

Abstract: As obesity becomes increasingly common worldwide [1], more people want to lose weight to improve their health and image. According to the Centers for Disease Control and Prevention (CDC), long-term changes in daily eating habits (such as regarding food/ nutrition type, calorie intake) are successful at keeping weights off [2]. Therefore, it would be helpful to have an artificial intelligence (AI) mobile program that identifies the types of food the user consumes and automatically calculates the total calories. This paper examines the development and optimization of an 11-categorical food classification model based on the Mobile-Net neural network using Python. Specifically, it classifies any food image as one of bread, dairy, dessert, egg product, fried food, meat, noodles, rice, seafood, soup, or fruit/vegetables. Methods of optimization include data preprocessing and learning rate and batch size adjustments. Experimental results show that scaling image inputs to standard size (Python Numpy resize) function), 300 training epochs, dynamic learning rate (start with 0.001 and *0.1 for every 30 epochs), and a batch size of 16 yields our best model of 83.44% accuracy.

Keywords: Food classification; Python; Data preprocessing; Mobile-net; Epochs; Overfitting; Learning rate; Batch size

Publication date: October 2021; ***Online publication:*** October 26, 2021

1. Introduction

With the modern society development, people's living standards are improving day by day, and the demand for dietary nutrition and health has also increased [3]. With the gradual popularization of scientific knowledge of dietary nutrition and health, they are gradually understood and pursued by the general public. On one hand, various nutrients required by the human body can be obtained through diet, such as protein, fat, vitamins, water and inorganic salts, etc. On the other hand, the food intake by the human body should be suitable for digestion and absorption, and the food should be fresh. Pollution and pollution-free, food processing should be scientific and reasonable, nutrition should be ensured, appetite should be improved, taste and flavor should be ensured, and calorie intake of three meals a day should also be properly allocated. Based on the suggestions provided by the nutrition associations, the energy distribution at 3 meals per day accounts for 30%, 40%, and 30% of the whole day respectively. Most people with obesity lack nutritional knowledge, and the male are suffering a more serious issue than the female, so the problem of dietary malnutrition is also more obvious [4].

According to the CDC [1], 42.4% of U.S. adults were obese in 2017-18, although the proportion was only 30.5% in 1999-2000 [2]. Notably, obesity is the most prevalent among people with lower educational levels and lower to middle income [5], who tend not to have the knowledge to design healthy diets or the money to hire nutritionists to do so [6].

1.1. Open Problem: The lack of nutrition knowledge and the challenge of knowing the complete knowledge base

Even though there has been an increased awareness of the nutrition health over the past years, the nutrition remains as a very specialized field, so that most people still cannot fully master the knowledge base or apply it correctly in the daily life. People might have read a number of articles, news and books on the healthy way for nutrition, but it is a challenge for everyone to consistently and correctly follow the best practice in any circumstances.

1.2. Solution: A mobile system to automate analyzing the food type and calorie using AI and deep learning

As a result, we developed a program to identify foods and nutrition types from photos and calculate the approximate calorie intake. Firstly, this would allow users to identify what kinds of food/nutrition they lack and adjust their diets accordingly. Secondly, estimating the user's calorie intake helps the user control the amount of food he/she eats and thus better stay on her diet to keep weights off. The program's primary features include identifying food types from user input images and estimating the total calories. Secondary features include graphing daily intake curves for intuitive understandings into how well the user is following his/her diet and a "share" button for users to share their meals.

The rest of the paper is organized as follows: Section 2 gives the details on the challenges that we met during the experiment and designing the sample; Section 3 focuses on the details of our solutions corresponding to the challenges that we mentioned in Section 2; Section 4 presents the relevant details about the experiment we did, following by presenting the educational impact in Section 5; Related work is discussed in Section 6. Finally, Section 7 gives the conclusion remarks, as well as pointing out the future work of this project.

2. Challenges

2.1. Challenge 1: Getting the complete and optimal dataset

Even though there are a large number of food images available online, the key challenge is the unavailability of the labelled images. All the experiments will be conducted using the needed label data for training come from in order to correctly classify and verify various foods. Generally speaking, in the downloaded data image training set, the number of different categories is unbalanced. In addition, the image does not always come in square shape, so a number of dataset pre-processing will be required.

2.2. Challenge 2: Conducting the experiment using the limited computational power

The fundamental quality of the deep learning application always lies in the quality and quantities of the dataset. As we are growing the number of available images in our dataset, it also increases the challenges of conducting the experiments with the limited computational power. How to balance the training quality and the training time becomes a new challenge in the field of deep learning. The best solution not only achieves the highest accuracy, but also can get the accuracy within a shorter period of training time.

2.3. Challenge 3: Selecting the correct and optimal deep learning models

When it comes to model selection and training tuning, challenges may arise as well. For example, if the model is too big, it will not be good for moving context. Besides, if the model itself may not be accurate. We have been seeing a growing number of deep learning models developed and turned from the academia in the recent years. Although the new models and examples provide a lot more options to tackle the proposed problem, it also generates new challenges to make the most optimal selection from the available algorithms. Generating the comprehensive accuracy result will take more efforts and the new models also required a bigger number of training images in the dataset.

3. Solution

3.1. Overview of the solution "healthy diet"

The app "Healthy Diet" is able to identify the food type from images, calculate food calories, allow the users to make healthy diet plans with existing ingredients at home, and realize the diet with healthy, attainable methods.

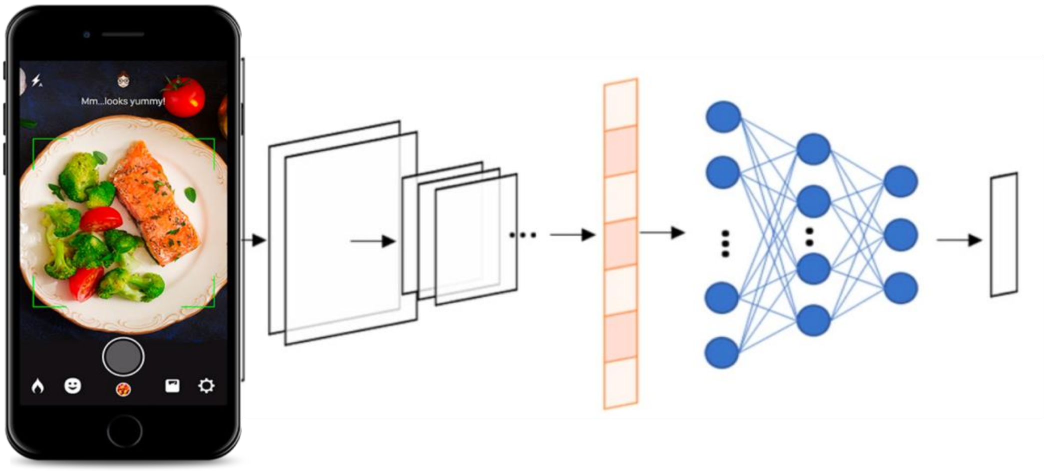

Figure 1. The deep learning-based solution to automate the food analysis

Figure 1 shows an overview of the solution. Health Diet can estimate the calories in the food. By having users taking pictures, Health Diet would identify the type of food through image recognition, retrieve the calorie data of the food from the database, and then use the approximate quantity of each food input by the user in order to multiply and estimate the calories in the food.

3.2. Features

In order to provide a decent user experience. A number of features have been proposed and implemented as follows:

(1) Meal Recommendations - By identifying all the ingredients, Healthy diet can provide users with healthy and delicious meal suggestions which match their taste. The app could guarantee a variety of healthy and low calories meals.

(2) Calculating Calories - Users can record their own diet, generate a growth chart to count their own calorie intake over a period of time, which allows Health Diet record, analyze, and promote health eating habits.

(3) Share Eating Habits - Users can share their own low-fat meals and track their own weight loss progress, which allow them to encourage each other to keep moving forward.

(4) Share Self-made Meal Plans - Users can upload their self-made dishes, and after passing the review, the meal plans will be added to the recommended menu on the app and shared with other users.

4. Experiment

The major contribution of the work lies in the deep learning model training and the selection of the most optimal model and parameters to tackle this problem. Some of the major research questions we are aiming to address are:

(1) The number of epochs (number of times the program loops through the dataset. Analogous the number of times a student reviews course materials) does not impact the size of the model.
(2) Decreased batch size increases model accuracy.
(3) Model accuracy initially increases as the number of epochs increases, but would start to decrease at some critical point. The rate of change is always negative.

4.1. Deep learning models

The neural network model has tried multiple layers of neural networks, convolutional neural networks, VGG, ResNet, EfficientNet, Mobile-net and other structures. Among them, EfficientNet did not train successfully, because the log cannot be seen to determine the cause; at the same time, Mobile-net has made progress, and the model recognition rate and size are acceptable, so I will focus on the optimization of Mobile-net later.

4.2. Experiment computing environment

The training machine is mainly carried out on Tencent's smart titanium server, and it also uses its own computer and Kaggle for code debugging and simple model training.

The initial multi-layer neural network and convolutional neural network were carried out on my own computer without a GPU. I remembered that training for 50 or 60 epochs would take a whole night to run. Starting from VGG, after adjusting the script, it will run on the Tencent Smart Titanium server.

4.3. Training dataset analysis

The training data set are the 11-image-dataset that was downloaded from Kaggle. It classifies food into 11 categories (**Figure 2**): "bread," "dairy products," "dessert," "egg," "fried food," "meat," "noodles/pasta noodles," "rice," "seafood," "soup," "vegetable/fruit."

As show in **Figure 3**, the compressed package size of the entire data set is 1.1G, divided into training/validation/evaluation three first-level sub-directories (number of files 9866/3430/3347), each sub-directory has a varying number of 11 food types second-level directories, in jpg format. The number of images for each food category is different. Taking training as an example, there are 1500 pictures in the Dessert/Soup secondary directory with the most, and 280 pictures in the Rice secondary directory with the least. The validation/evaluation are similar, the number of pictures of each food is less, and the ratio is roughly the same.

Bread 994	Bread 362	Bread 368
Dairy product 429	Dairy product 144	Dairy product 148
Dessert 1500	Dessert 500	Dessert 500
Egg 986	Egg 327	Egg 335
Fried food 848	Fried food 326	Fried food 287
Meat 1325	Meat 449	Meat 432
Noodles-Pasta 440	Noodles-Pasta 147	Noodles-Pasta 147
Rice 280	Rice 96	Rice 96
Seafood 855	Seafood 347	Seafood 303
Soup 1500	Soup 500	Soup 500
Vegetable-Fruit 709	Vegetable-Fruit 232	Vegetable-Fruit 231
totally 11 kinds, 9866 training pictures.	totally 11 kinds, 3430 validation pictures.	totally 11 kinds, 3347 evaluation pictures.

Figure 2. Training data set of food

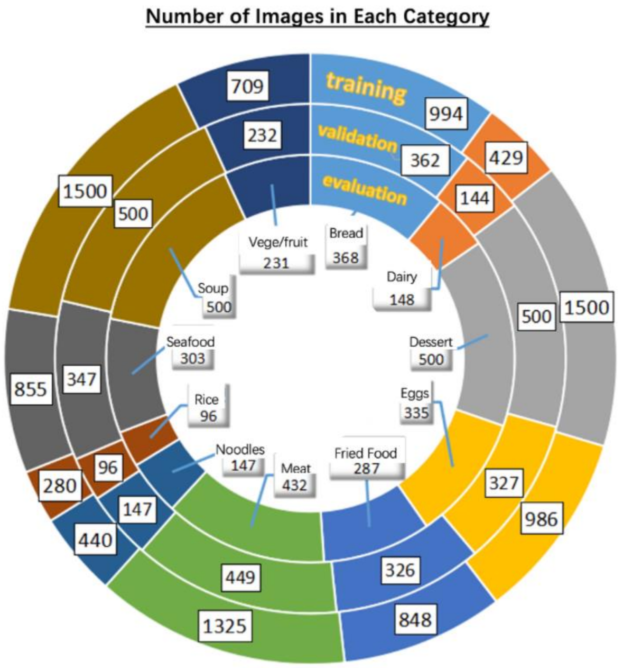

Figure 3. The distribution of the training data set category

If the training data is extremely unbalanced, image preprocessing, under-sampling or oversampling or supplementation is often required, in order to prevent the model results from being biased towards the category of large data.” There happened to be a period of test data comparison, but from the results, it is not very consistent. The test model uses 100 rounds of training results of ResNet18, and 15 pictures with known correct classifications are taken from each of the 11 types of food for prediction verification. The test result of a total of 165 pictures: the most soup, 15 pairs of all pairs are expected. The accuracy rate is 100%; for the same number of desserts, only 7 predictions are correct, which is the bottom of the accuracy rate as fried food, less than 50%; with so few pictures of rice, it predicted 9 pictures correctly (60%), which is slightly lower than the overall accuracy of 67.88%. Based on our experiment, we believe the cause for this issue are:

(1) The knowledge seen on the Internet is for reference and could contain some hidden parameter variance.

(2) Its conclusions often require certain preconditions (e.g., whether our data should be classified as “extremely unbalanced training data”), and there are also models that are not up-to date.

(3) The problem can also be resulted from the issues such as insufficient training and insufficient test sample size.

This is a research question that we need to verify in the process of learning and practice.

We have manually browsed and checked the catalogs. Most of the pictures (more than 90%) are square (the same number of pixels in length and width), a few (3-5%) are rectangles with different lengths and widths, and the few are the least. Exaggerated the author has seen graphics with an aspect ratio close to 2:1 (for example, 512*288). The narrowest (length or width) pixels seen have 280+ pixels. In fact, this part of the data characteristics was not deliberately paid at the beginning. This was only noticed when thinking about how to improve the accuracy of the model in the later stage. Because improving the quality of training images is also an important means to improve accuracy.

After reviewing what has been done in some other related work [5][6][7], the Mobile-net model can reach 90%+ after various skills blessings, but we only have 70%+. This issue will be discussed later. Then one of the concerns is the training image. For example, it's too late to do high-tech physical work like checking and correcting images one by one. First, we cast my suspicion on resize. When we specify the length * width parameters when training the data, the machine will resize, so what about this resize? Is it cut or scaled? A search on the Internet should be the latter, which is also in line with our thinking and judgment-if it is cutting, the machine needs to have a certain amount of intelligence to judge the choice of images, but this is precisely what we want to train through the AI model to let the machine learn the skill, it is unlikely to master this skill in advance. Since it is zooming, there is a deformation problem, so you need to check the image, visually, more than 90% are (or close to) square, then it is almost the same, this factor should not have much effect.

Another factor is that when resizing, it is best to reduce the original image-not to enlarge it. Visual inspection-can only rely on visual inspection first-the smallest length/width pixel value is also close to 300, so resize to 224*224. It shouldn't be a big problem.

In terms of input images, if the plan can be improved without time for implementation, at least some images with a large aspect ratio can be deleted under a large number of food types (such as desserts and soups).

4.4. Experiment results

Figure 4 shows the overall performance of the selected machine learning models. The detailed results and comparison will be discussed in the following sub sections.

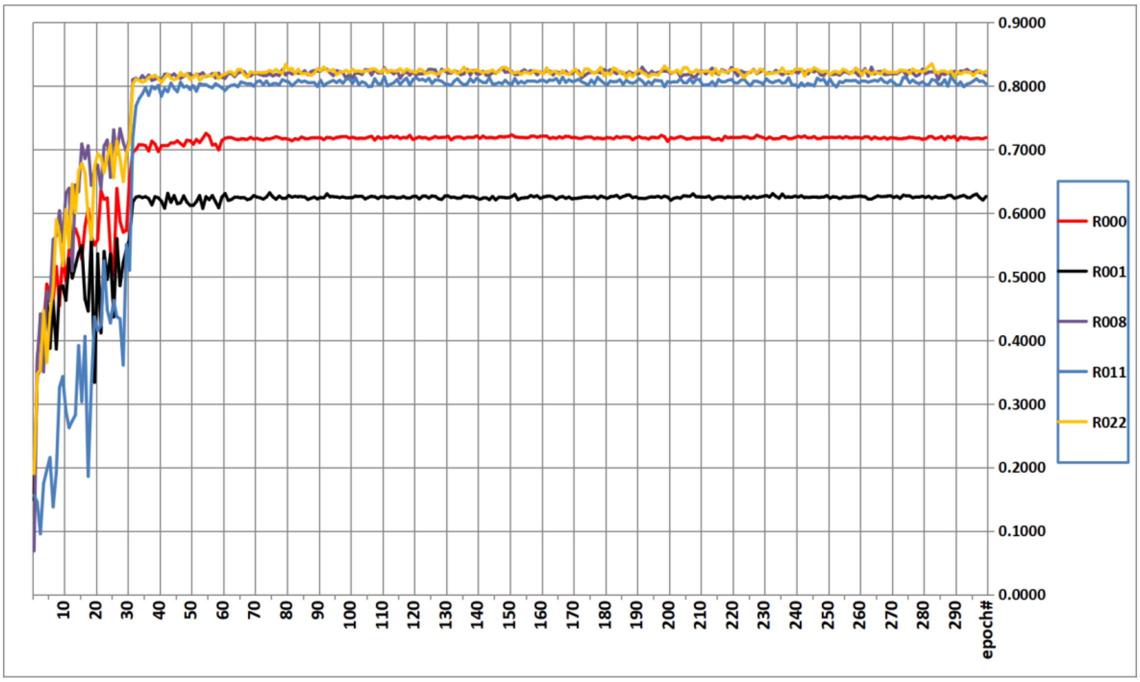

Figure 4. The Summary of the Machine Learning Model Performance

4.4.1. The performance comparison using MLP, CNN and VGG

This stage is in the running-in period of environment construction and code configuration with setup. Among these models, MLP and CNN [7-8] can still be trained on their own environment, and it takes about ten hours to train 60 epochs. It takes more than 8 hours to train 3 rounds of VGG, and it can only be trained on Tencent's smart titanium server, but the initial period may be improper use of the method, and no log can be seen. The log problem was solved in the later stage of ResNet and after Mobile-net.

The problem encountered in the experiment is that after more than 30 epochs of CNN training, the training accuracy gradually increased to 95%+, while val_accuracy did not increase after 0.3x but slightly decreased. Checking the information relevant and thinking together, this should be an over-fitting phenomenon. After adding the Dropout() layer to the model, the phenomenon is alleviated, that is, val_accuracy no longer shows a downward trend, but after rising to a certain value (0.5), the small-range fluctuations no longer rise.

The corresponding VGG model can run normally on the smart titanium server, because the log feedback cannot be seen, and the training accuracy and val_accuracy data during the training process cannot be known. It is estimated that its performance should be better than CNN.

For the above three models, the results are unsatisfactory when the images in the evaluation are called by themselves. The h5 models are all over 200M in size, and they are all too large to be used in small mobile programs. It must be studied forward.

4.4.2. The performance analysis on ResNet model

We tried ResNet18 [9], by following the given reference code and library documentations, and the initially tried image enhancement techniques, including horizontal flip, up and down translation, or left and right translation. Dynamic learning rate is used for initial learning.

The learning rate uses the initial value of 0.001 (or 0.01/0.0001, etc.), and shifts one decimal point to the right every 30 epochs.

The size of the h5 model is reduced to 135M. It’s still too big. The test recognition rate reached 68%, and it was initially available.

According to thinking, the image enhancement technology in the reference sample code is moved from val_data_gen to train_data_gen. Because image enhancement should be used in the training session, which is equivalent to the effect of increasing the number of training images; and it should be relatively less important when applied to the verification link.

Based on the reference document, it says that by reducing the number of down sampling (that is, when certain links are convolved, the original model stripe=2 is changed to stripe=1), it improves the model recognition rate. Taking into account that the training data used by the teacher for lectures is 32*32 numbers, it is necessary to reduce down sampling to avoid the image falling too early to 1*1, and our model is 224*224, which is the standard input of ImageNet, which may not be necessary. Therefore, according to the teacher’s courseware and the standard map on the Internet, the teacher’s two down sampling codes were found, and the standard ResNet18 model was restored.

We have tried to learn the ResNet50 structure by ourselves and try to rewrite it on the basis of the ResNet18 code provided, but it was not very successful. At the same time, we tried to modify the code of EfficientNet, but failed to generate the model. Thus, due to the successful debugging of Mobile-net, the author concentrated on the research and development of Mobile-net

4.4.3. The performance analysis on Mobile-net model

After the model was commissioned, the training was continued in units of 10 epochs. According to the log judgment, the optimal val_accuracy appeared at epoch=55, which was 72.52%, but at that time, the training was continued and saved once every 10 epochs, so the model was not saved. The model that has a chance to be saved in the future is 72.05% of epoch=218.

The size of the Mobile-net [10-11] model is about 1/9 of the ResNet model. The h5 model is 15M, after the conversion is 6.8M, it is relatively faster to loAad on the mobile phone applet.

As the research on other parameters of Mobile-net, at the beginning, it was continuously submitted in segments, and 300 epochs were trained successively, hoping to find a better val_accuracy value. The exploration of other parameters only started these days. However, the response of the smart titanium server seems to have changed. In the past few days, it has been found that val_accuracy has generally decreased, including re-running the original 300 training scripts, which is even difficult to regain 65%. Explain this re-run. Strictly speaking, it is not to rerun the original script verbatim: it turned out to be a script submitted every 10 epochs after the call, and then collect the logs and submit the next one and a half hours later. For scripts, submit a training script for 50/60 times at one time during sleeping hours or long-term classes when it is inconvenient to surf the Internet; now it is changed to submit a training script for 300 epochs at a time, and it is automatically saved in the middle. As the original grading script is very regular, other parameters have not changed except epoch, so the one-time 300 epochs submitted now, self-understanding is an equivalent python script. However as a beginner, the author might have overlooked something, so the first thing the author can't rule out is my own reason (although the author couldn't find this reason). Secondly, during the study and practice during this period of time, the author realized that a large system, such as Tencent Smart Titanium Server Cloud, has many people working hard and many factors affecting its operation. It is not ruled out. The combined effect of several factors happened to affect the environment of this AI training. For example, the random number generation sequence happened to have a certain change. In short, it boils down to one sentence: In recent days, it was discovered that the val_accuracy (this is the main evaluation index we found) of the submitted mobile-net training scripts has generally declined. The reason is unknown, which makes the trial effect of some new parameters unable to be evaluated.

Therefore, the points mentioned later are all ideas and possible means. To evaluate whether it is useful, just look at the results of a two-round running script, there should be a lot of chance. It may take a longer time, fix other factors each time, and specifically study and compare the influence of a certain factor change, and the conclusion will be more reliable.

In order to obtain the historical peak value of val_accuracy (72.52% at epoch=55), we tried to start with the saved epoch=50 model and re-run 51-60, saving the model at every step. As the image enhancement of the training parameters is inherently random, the parameters must be different during the actual operation. The author has tried two scripts, one is that the log method is not changed at all (see the description of simplified log below), and the other is that the log writing is theoretically irrelevant to model training and we only care about val_accuracy, so the log is simplified. The result is that the highest version of the unmodified log has reached 71.61%, 71.55%, 71.41%, etc., and the simplified version of the log has 71.14% twice. However, the previous (referring to the 1-300#epoch run in multiple runs) has reached more than 71.8%, but the new 51-60#epoch in these two rounds has not reached. This test was only remembered today, and it was done temporarily. It may have a certain supporting role in supporting the "small background situation" mentioned in the previous paragraph.

In the past few days, we have also tried to add image enhancement parameters to val_data_gen, but the result is that val_accuracy has decreased (peak value is more than 60). The author is not sure if it is the reason for the general decline in val_accuracy in the last few days or the decline caused by stricter verification. According to thinking, the validation link and training use different pictures independently, so this change should not affect the improvement of the training effect itself, but affect the consistent standard of evaluation. Therefore, in other attempts, no image enhancement parameters were applied to val_data_gen. The image enhancement in the train link adds the rotation_range=20 parameter, which is intended to randomly rotate the image by plus or minus 20 degrees. In effect, val_accuracy also drops, with a peak value of more than 60.

Similar to before, the live broadcast class teacher still used the code to reduce downsampling (3 places in total). We tried to add 1 additional drop and restore the 3 drop to change back to the standard model. We want to compare these two attempts. As a result, val_accuracy is still declining, and the peak value is still more than 60 points.

We also tried to simplify the log. By setting the environment variable TF_CPP_MIN_LOG_LEVEL=3 and the fit_generator() parameter verbose=2, only the val_accuracy we care about is displayed. The effect is good, and the log is much more concise.

Finally, the parameter workers=n in fit_generator() allows the maximum number of threads. Appropriate settings can increase the training speed. But in the process of gradually trying to increase this parameter, once workers=8 resulted in no log (it should be caused by other factors on the server. Later, re-run the script without changing the workers parameter and you can see the log again. This is only the truth. Record the test phenomenon).

5. Educational impact

When it comes to issues in K-12 educational settings, healthy diet can play an important role as well. According to Ruiz et al. [12], the children and adolescents with obesity has risen three times since 1970s, and those with server obesity has four times increment. It is also shown that children and adolescents with obesity are very likely to stay obese in their adulthood. It is an essential issue to treat obesity before adolescents enter adulthood, as obesity increase the likelihood to develop cardiovascular and mental illness. Healthy diet is widely applicable to this issue. If healthy diet is accepted by users in K-12 schooling, they will be able to create healthy meals on their own and learn how to keep their weight in a healthy range effortless. They do not have to worry about looking up random menus online or purchasing specific ingredients for a special diet. healthy diet is able to plan a health meal with the existing ingredients back at home, and it keeps tract of the calorie intake. Students can evaluate their progress just by accessing healthy diet.

On top of that, besides treating the obesity, Healthy diet is also capable of ameliorating eating disorder. Social media has created unrealistic body images for adolescents, which has greatly jeopardized their mental health. It has already caused eating disorders to be a ubiquitous problem among teenagers. Adolescents are often eager to become skinnier, so they stop having any food several days in a row. Not only does this behavior harm physical health significantly, but it also causes body weight to fluctuate even more in the long term. If healthy diet is by adolescents, they will be able to lose their weight in a healthy matter, and maintain healthy lifestyle meanwhile. They can also evaluate their calories intake through viewing the sample menu to see if they are on the health level.

6. Related work

Related works solve some problems such as inaccuracy, volume of food, and inability to recognize all food in an image [13]. Manika Puri, Zhiwei Zhu, Qian Yu, Ajay Divakaran, and Harpreet Sawhney [14] found a solution to old food intake assessment that suffer from inaccuracy or complex lab measurements. Their solution is to use a mobile phone to capture images of foods, recognize food types, estimate their respective volumes and finally return quantitative nutritional information. Yuji Matsuda, Hajime Hoashi, and Keiji Yanai [15] proposed a two-step method to recognize all the food in multiple-food images. Unlike these related works, we propose to apply these methods into developing a solution that helps people log their food and calories and ultimately helping them form a habit of recording their meals.

Liu et al. [16] applied an edge computing technique to apply deep learning on food recognition. The major difference between their work and our is that they focus on identifying the food type only, while we are also analyzing the calorie amount. Similarly, Pouladzadeh et al. also applied the deep learning technique in the same domain, but their work targeted on trying to identifying multiple food images in the same image [17].

7. Conclusion and future work

In this paper, we have presented a mobile application system to automate the analysis of food calories using deep learning and neutral network. According to the experiments we have conducted, the standard Mobile-net can achieve a recognition accuracy of 76.69%, and after adding the two technologies of “reduction down sampling” and “data augmentation,” it can achieve an astonishing 92.10% accuracy. For the food 11 Mobile-net model we trained here, the data augmentation has been included. The reduction of down sampling is modified on the basis of the reference code, and 3 reduction measures have been passively introduced. Thus, our Mobile-net model should be against 92%, but it is actually only 72%. We think that the live broadcast class teacher uses the cifar10 data set, which recognizes handwritten numbers. Compared with the current food category, the difficulty should be different, so the decline in recognition accuracy should be reasonable.

In terms of the number of the users, the later model training, many explorations and parameter adjustments only improved the accuracy of a few percentage points or a few tenths of a percentage point. In this demonstration, increasing the accuracy of these percentages may have little effect. However, in many projects with a huge user base (possibly hundreds of millions) such as Tencent, this gap of a few thousandths or a few ten thousandths will have a huge impact. Therefore, whether some improvement methods are worth the effort depends on the environment in which they are placed, the user base and the accuracy requirements. The experiment conducted has been very rewarding. This prototype has systematically sorted out the knowledge system of mathematics and artificial intelligence for us, and has opened the door for more foundation in AI. Although there are a lot of confusions in the learning process, we have gained more. It lays the foundation for our further study and understands a system. When we have the opportunity to learn relevant knowledge systematically in the future, we will know the direction of learning better.
Regarding the future work, we will be mainly focusing on two major aspects:

(1) How to apply reinforcement learning in this problem and verify its accuracy.
(2) Performing more experiments with the increased number of training dataset and check the influence of the training dataset number.
(3) We also want to build a mobile application that allows users to simply take a picture and get the result promptly.

Disclosure statement

The author declares no conflict of interest.

References

[1] Bray GA., ed., 1979, Obesity in America: A Conference.

[2] Oliver J, Taeku Lee, 2005, Public Opinion and the Politics of Obesity in America. Journal of Health Politics, Policy and Law 30, 5: 923-954.

[3] Bean M, Stewart K, Olbrisch ME, 2008, Obesity in America: Implications for Clinical and Health Psychologists. Journal of Clinical Psychology in Medical Settings 15(3): 214-224.

[4] Hales CM, Carroll MD, Fryar CD, Ogden CL, 2020, Prevalence of Obesity and Severe Obesity Among Adults: United States, 2017–2018. NCHS Data Brief, no 360. Hyattsville, MD: National Center for Health Statistics. [https://www.cdc.gov/nchs/products/databriefs/db360.htm]

[5] 2021, Centers for Disease Control and Prevention. (2020, August 17). Losing Weight. https://www.cdc.gov/healthyweight/losing_weight/index.html.

[6] Eagle T, Sheetz F, Gurm R, et al., 2012, Understanding Childhood Obesity in America: Linkages Between Household Income, Community Resources, and Children’s Behaviors. American heart journal 163(5): 836-843.

[7] Lei X, Pan H, Huang X, 2019, A Dilated CNN Model for Image Classification. IEEE Access 7: 124087-124095.

[8] Liu Z, Wei QY, Mee LY, 2018, Image Denoising Based on a CNN Model. In 2018 4th International Conference on Control, Automation and Robotics (ICCAR), IEEE, 389-393.

[9] Bethge J, Bartz C, et al., 2020, MeliusNet: Can Binary Neural Networks Achieve Mobilenet-Level Accuracy?. arXiv preprint arXiv:2001.05936.

[10] Qin Z, Zhang Z, et al., 2018, Fd-mobilenet: Improved Mobilenet with a Fast Downsampling Strategy. In 2018 25th IEEE International Conference on Image Processing (ICIP), IEEE, 1363-1367.

[11] Wang W, Li Y, Zou T, et al., 2020, A Novel Image Classification Approach via Dense-MobileNet Models. Mobile Information Systems 2020.

[12] Ruiz LD, Zuelch ML, et al., 2019, Adolescent Obesity: Diet QUALITY, Psychosocial Health, and Cardiometabolic Risk Factors. Nutrients, 12(1): 43. https://doi.org/10.3390/nu12010043.

[13] Kawano Y, Yanai K, 2013, Real-Time Mobile Food Recognition System. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 1-7.

[14] Puri M, Zhu Z, et al., 2009, Recognition and Volume Estimation of Food Intake Using a Mobile Device. In 2009 Workshop on Applications of Computer Vision (WACV), IEEE: 1-8.

[15] Matsuda Y, Hoashi H, Yanai K, 2012, Recognition of Multiple-Food Images by Detecting Candidate Regions. In 2012 IEEE International Conference on Multimedia and Expo, IEEE: 25-30.

[16] Liu C, Cao Y, Luo Y, et al., 2017, A New Deep Learning-Based Food Recognition System for Dietary Assessment on an Edge Computing Service Infrastructure. IEEE Transactions on Services Computing, 11(2): 249-261.

[17] Pouladzadeh P, Shirmohammadi S, 2017, Mobile Multi-Food Recognition Using Deep Learning. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) 13(3s): 1-21.