Int. J. of Electronics Engineering and Applications, Vol. 9, No. 3, July-Sep. 2021

FOOD-IMAGE CLASSIFICATION USING NEURAL NETWORK MODEL

Alex M. Goh and Xiaoyu L. Yann

ABSTRACT

New digital invented technologies are changing almost all the industries in the world. Food and beverage industries are one of them. Hotels, resorts and restaurants are using their good-looking images to attracts the customers. In the other side, customers are unaware about the originality of the images. To classify food images, machine learning may be a good solution. In this proposed work we have used a convolutional neural network for food-image classifications. As a convolutional neural network removes spatial features from images, so it is very efficient for food-image classifications. This proposed model may help restaurants owners for advertisement of their food to people looking for the same type of food they offer. Additionally, this model may be used for the food a distribution system. We have developed a neural network model to classify the food- image. We have also used the transfer learning technique with Inception V3.

Index Terms: Machine Learning, Food-image, Data Augmentation, Convolutional Neural Network, Transfer learning, Inception-v3.

Reference to this paper should be made as follows:

Alex M. Goh and Xiaoyu L. Yann, (2021), "FOOD-IMAGE CLASSIFICATION USING NEURAL NETWORK MODEL" Int. J. of Electronics Engineering and Applications, Vol. 9, No. 3, pp. 12-22, DOI 10.30696/IJEEA.IX.III.2021.12-22

Biographical notes:

Alex M. Goh *was born in Malaysia, completed his Bachelor and Master degree from Multimedia University, Malaysia and Ph.D. from Curtin University, Malaysia, currently he is working as Data Scientist in Multimedia University, Multimedia University.*

Xiaoyu L. Yann *was born in China, He completed the bachelor's and master's degree from Hebei Normal University. He is currently pursuing PhD with Curtin University, Malaysia, His research interest includes Artificial Intelligence and power grid operation analysis.*

1. INTRODUCTION

With the advent of social media, thousands of photos are uploaded to social media like in Facebook, Instagram, Telegram etc. In the other part, the rise of mobile devices puts cameras in everyone's pocket. Active-users were no longer able to log into their social media accounts without having a desktop device and it started quickly. Instagram-based social networks such as Instagram and Snapchat have emerged to live up to this new reality on the same phone or tablet you use to engage and take photos on social media platforms on the go. Twitter responds to user behavior by giving you more opportunities to center your images in front. Marketers know how important social media marketing is.

The information that overflows on the Internet, social media platforms, is enormous. This data represents the challenges and opportunities we strive to market effectively, protect our image, and enter the era of information overload [1]. The potential hidden in this constantly growing pool of online images. For the brand, this means accessing more data than ever before, especially image-based data. Social media users fully embrace the concept of sharing photos instead of text or text. Snapchat has 184 million users per day[2]. In 2017 (up from 46 million in early 2014), Instagram increased from 800 million in September 2017 to 1 billion monthly users identifies this growth trend the ability to analyze, analyze and utilize image recognition technology had to exist in the future. Most of his digital marketing was dominated by visual data. Otherwise, the brand would have missed an entire pile of valuable data. Media monitoring may not be captured. If you miss a great opportunity to learn and communicate with your customers, artificial intelligence and image recognition make it easier for marketers to find visuals on social media without explicit textual mention [3,4]. Food images are being uploaded on social media and with food identification, social media can group people based on their food choices. Social media platform for advertising target audiences. Computer vision and image processing techniques are currently being used in many fields. Image: Food identification is a challenging task because food images have less variation within the classroom. Data Based Classification It is possible to use it in a cheaper device. Today, there are inexpensive smartphones with high computing power that are capable of processing high-definition image data, so the model described in this article can be made on smartphones.

Machine learning a new branch of research has given suitable platform for the above discuss domains [5,6,7]. Deep learning which is considered as sub-set of the machine learning is more specific for handling social media problems [8,9]. Transfer learning may be defined as a sub-branch of the deep learning. Transfer learning is the reuse of a pre-trained model for a new problem, it is very popular nowadays in deep learning because it can train deep neural networks with relatively little data, and it is very useful in data science because of most real problems. You don't have millions of data points marked to train these complex models [10,11]. Let's take a look at what transfer learning is, how it works, why and when to use it. Includes several resources for models that have been previously trained in learning transfers for example, when you train the classifier to predict whether an image contains food, you can use the knowledge gained during training to recognize drinks, for example, if you trained a simple classifier to predict, if the image includes a backpack, you can use the knowledge gained by the model during training. Recognizes other objects such as sunglasses. In passing on knowledge, we try to basically apply what we have learned in one task to improve the generalization in another. We transfer the weights that the network learned in "task A" to the new "task B". The idea is to use models that have learned from business in a new business with a lot of data training cards available and with little data. Instead of starting with learning processes from the beginning [12,13].

Transfer learning is mainly used for natural language processing tasks such as computer vision and emotion analysis due to the large amount of computing power required. Transfer learning is actually a machine learning technology. No, but we can think of it as. For example, design methodology in areas such as active learning. It is not the exclusive part or research area of machine learning. Nevertheless, it is very popular in combination with neural networks that require large amounts of data and processing power [14,15].

We may use the different datasets for the food categories and food-101 dataset with 101 food categories is one of them. All images are rescaled to a maximum side length of 512 pixels. Use a subset of the four food categories [Fry-Chicken, Mutton, Bread, Omelet, Waffle]. See Figure 1 for this assignment. The data consists of three main subfolders: training, validation, and testing. The training data consists of 1000 images per class, with up to 500 validation images and up to 500 test images per class. The dataset has not been (intentionally) cleaned up and therefore contains [16,17]

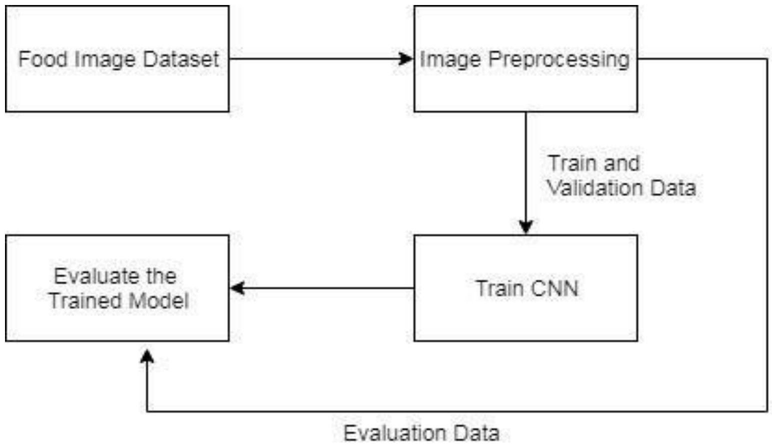


Fig1. The flow chart of multiclass food classifier.

2. METHODS

A. Data Augmentation

Data augmentation is an important part of Food image classification. The appropriateness of data augmentation will directly affect the final effect of food image classification. Food the labeling of these data can only be done by professional radiologists, and it needs complex experienced radiologists to spend half an hour or more time to observe image over and over again and hand out the location of the nodules. As a result, medical image data inevitably suffers from a lack of data due to the scarcity of professionals (far fewer than the number of professionals who can annotate other voice or text data). The significance of data augmentation lies in that, on the limited Food image data set, the data amount of training can be expanded sufficiently and reasonably to improve the generalization ability of the model. Reasonable and appropriate noise data are added to improve the robustness of the model. In order to eliminate dimensional effects between data.

ImageNet image classification contest with an accuracy rate of 10% over the second place. Since then, the convolutional neural network model has returned to the vision of researchers. VGG model , Google Net model , ResNet model , NASNet model and other network models have sprung up.

In general, many independent neurons can form a two- dimensional plane, and the deep convolutional neural network is composed of many layers in which many two-dimensional planes of feature mapping form. A deep convolutional neural network consists of four core parts.

The first part is the local perception which all neurons in the neural network do not need to perceive the global image, but only the local information, and global information is obtained by gathering local information.

The second one is the convolution. The function of convolution is to extract image features, and the number of parameters can be greatly reduced by using the convolution kernel.

The third is weight sharing. Weight sharing means that the parameters in the same convolution kernel are used for the whole picture, and the weight in the convolution kernel will not be changed due to the different positions in the image. Moreover, weight sharing of convolution operation can image data. (d)The fourth part is pooling. The pooling layer is usually placed behind the convolution layer in the convolutional neural network, which can be used to reduce the characteristic dimension of the output of the convolution layer of the previous layer, but at the same time retain the effective key information of image data.

B. Deep Convolution Neural Network (Dcnn):

The development of the convolutional neural network model is not plain sailing. 1980, Fukushima [14] proposed neurocognitive based on visual pattern recognition mechanism widely considered the first version of convolutional neural networks. It decomposes and com- bines the visual system layer by layer to model it so that it can correctly identify objects with displacement or slight deformation. In 1989, LeCun et al. proposed a convolutional neural network with a five-layer structure, named LeNet, which solved the problem of handwritten numeral recognition. It was the beginning of convolutional neural network from research to application. However, due to the lack of training data and the weak computing ability of computer hardware, the convolutional neural network was not developed that time. On the contrary, manual feature extraction with SVM performed well on small data sets leading to its mainstream research at that time. Then in 2012, AlexNet, a famous neural network model established by Krizhevsky et al., won the first place in the some noise. It is mainly displayed in dark colors and in some cases has the wrong label. The dataset is not complete, which makes the problem even more difficult. However, it uses the assigned label).We have taken some image pre-processing technique to increase efficiency to our system. First, we re-sized all our images to 224 x 224 x 3 to increase processing time and also to fit in our convolutional neural network model.

In this paper, Deep Convolutional Neural Network (DCNN) is used for extracting features. The reason why DCNN model is used for extracting features is that it has two characteristics: local connectivity and weight sharing. These two features are very useful for feature extraction in food images, which can avoid a large number of human resources input of repeated markers, and it can also ensure that the learned convolution kernel has a strong response to local features. Fig: Shows the process of the model

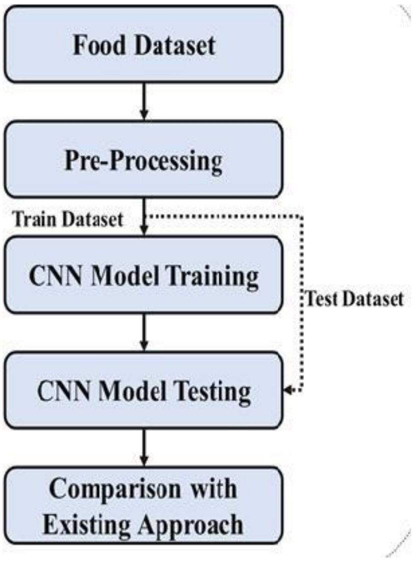


Fig2. Process followed by model

C. Transfer Learning

With the increase of the depth of the neural network, the performance of the neural network will indeed be improved, but this improvement is at the expense of time and computing resources. In order to reduce the cost of training, deep learning based on transfer learning emerges. In a word, transfer learning is to transfer the trained model parameters to the new model to help the new model training. Considering that some data or tasks are relevant, the learning efficiency of the new model can be accelerated and optimized by sharing the parameters of the trained model to the new model through transfer learning, instead of starting from scratch as most networks do. In recent years, due to the popularity of deep learning, the demand for data also increases, so transfer learning has been further developed. In order to train a high-quality classifier from the data of different distributions, Dai et al. proposed an instance-based AdaBoost transfer learning algorithm, which uses the training data to select the useful parts from the different distributions data, so as to reduce the classification error rate. In order to solve the problem of domain Adaptation in transfer learning, Long and Wang proposes Deep Adaptation Network, which extends Deep convolutional neural Network to domain Adaptation; Cao et al. proposed Partial Transfer.

3. EXPERIMENT

A. Data Description

In order to compare and analyze the results of other researchers on the classification of food images, the standard public digital image database food 101, Technology, was selected as the experimental data set in this paper. All images are rescaled to have a max side length of 512 pixels. We will use a subset of 4 food categories [chicken curry, hamburger, omelet and waffles]. Data is organized in three main sub-folders: train, validation and test. Our training data consist of 1000 images for each of the classes, we have ~500 validation images per class and ~500 test images per class. The dataset was not cleaned (on purpose), and thus still contain some amount of noise. This comes mostly in the form of intense colors and sometimes the wrong labels. The neural network

model based on transfer learning needs to take into account the size of the new data set and the original data set and the similarity between the two data sets.
Learning based on selective adversarial networks (SAN). It only transferred the partial samples related to the source domain and the target domain, and then dealt with partial transfer problems through SAN. Because the adversarial network can learn domain invariant features well, it can play a great role in transfer learning. Busto and Gall proposed Open Set Domain Adaptation. They use the relationship between the source domain and the target domain to label the samples of the target domain, and convert the source domain to the same space as the target domain, and finally classify the samples of the target domain.
The number of food images in the Food 101 database used in this paper , which is seriously insufficient com- pared with the amount of data required by the neural network model, so this paper adopts the method of transfer learning. Considering the performance of the model and computing resources, we chose Inception-v3 model as transfer learning framework which is trained on ImageNet datasets (including more than 1 million copies a total of 1000 categories of image data) and has a good performance on small data set. This paper refers to the structure of Inception-v3 of Google Net and makes some improvements because of its structure. In order to make the model more suitable for our experiment, the structure of the model is fine-tuned that the last three layers of the model are cut out, and the results of bottleneck layer are taken as the feature results. The structural schematic diagram of the Inception-v3 model.

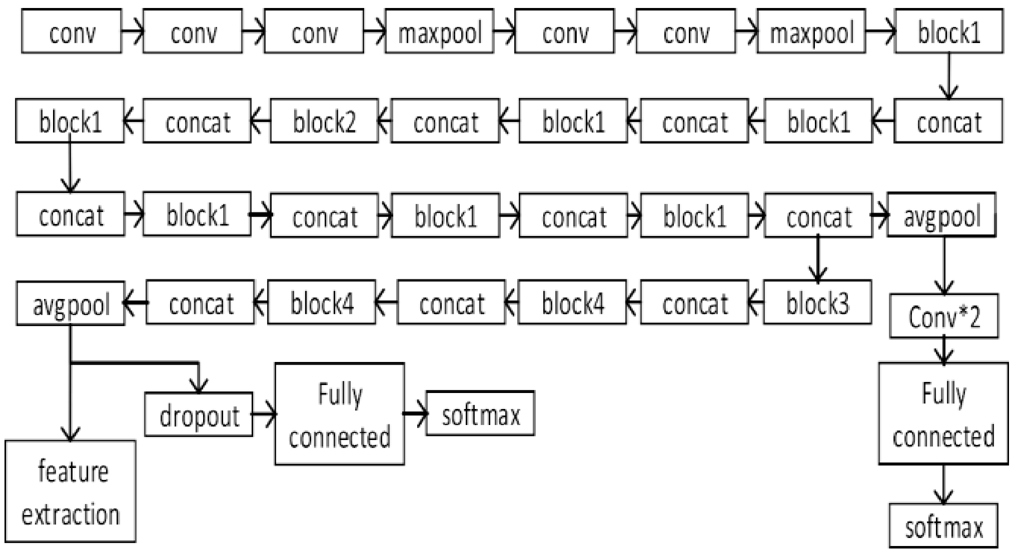


Fig3. The structural schematic diagram of the Inception-v3 model.

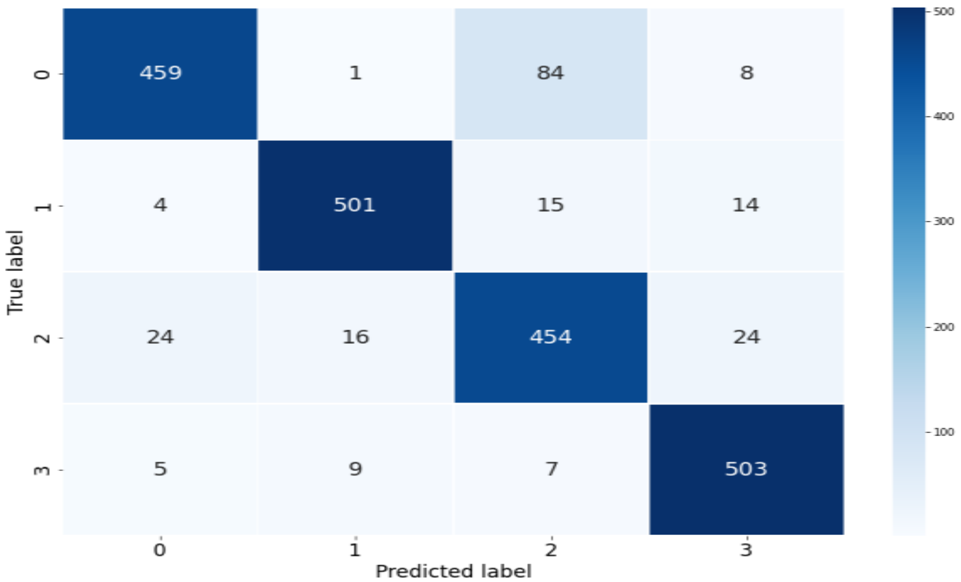
A. The Method of Evaluation

This paper uses the method of ten-fold cross-validation to verify the experimental results. Ten-fold cross-validation is to divide the data set into ten parts at the same size, take one part in sequence each time as the test set, the rest as the training set, and finally take the average of ten results as the estimate of the metrics of the model. The steps of ten-fold cross-validation:

a) Dividing the dataset into ten parts on average. Each part contains an equal number of images.
b) For each training, one part was selected in sequence as the test set, and the remaining dataset was used as model input to train the model.
c) After training, true-positive, false-positive, true- negative, false-negative of the test set were recorded successively.
d) Averaging the results of ten records to obtain the average sensitivity and average specificity of the final test results.

B. Result Analysis and Discussion

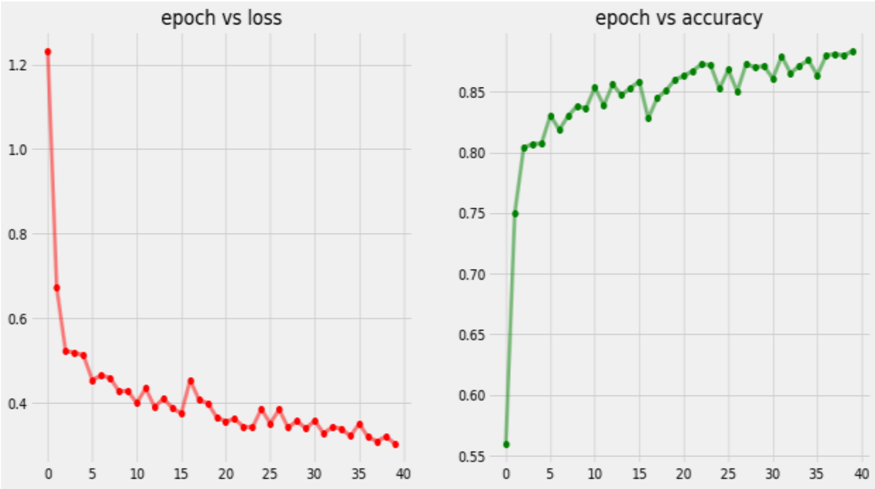
However, DCNN can automatically extract features so that it can save a lot of time and labor. Moreover, the sensitivity and specificity of SoftMax, Logistic and SVM connecting the fine-tuning Inception-v3 model increased by about 2% compared with the experimental results of DCNN. Although the Inception-v3 model was trained on the ImageNet dataset, compared with the DCNN model without transfer learning, its average accuracy has been further improved and ran faster. Experiments that fine-tune the Inception-v3 model connected different classifiers have resulted in significantly better results. The classification errors mainly lie in:
For migrating large-scale neural networks such as inception-v3, Because the network is only fine-tuned, the extracted features are not enough, which leads to some image classification errors.
Only three classifiers are selected to compare the result in this paper, and the selection of classifiers will affect the classification.
It can be seen that compared with other non-hand-crafted feature extraction methods, DCNN model has higher specificity and sensitivity, while its effect is lower than some hand-crafted feature extraction methods.
The Confusion matrix graph of model before tunning is:

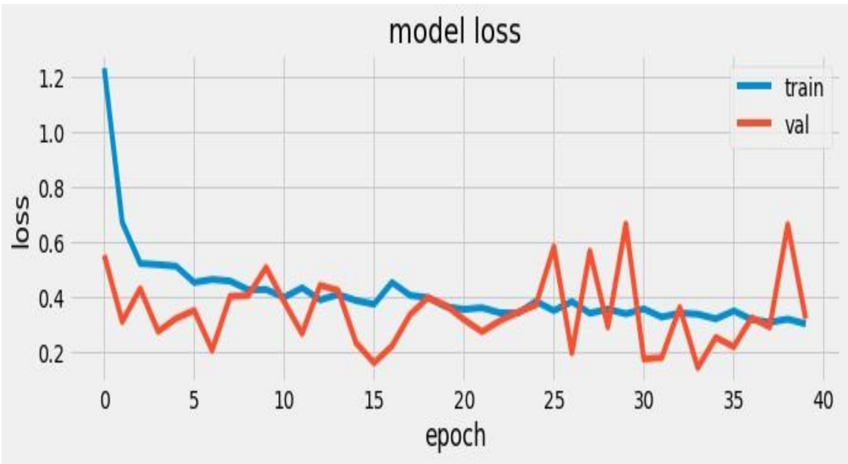


The Confusion matrix graph of model after tunning is:

1. Model Before Tunning

The graphs show the results before tunning the model.

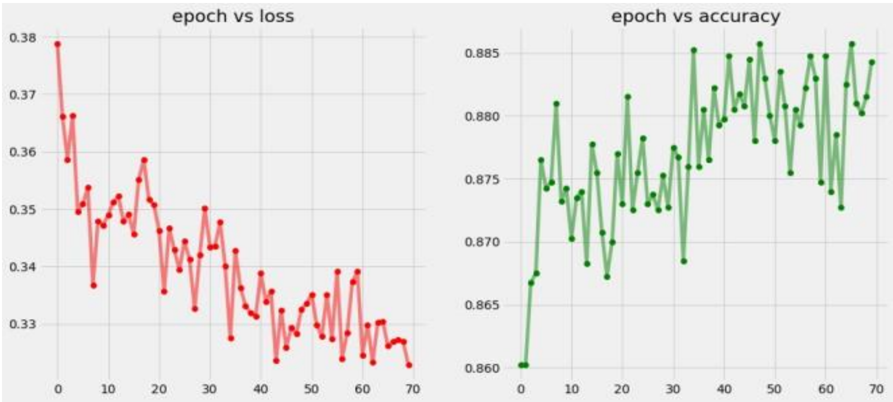


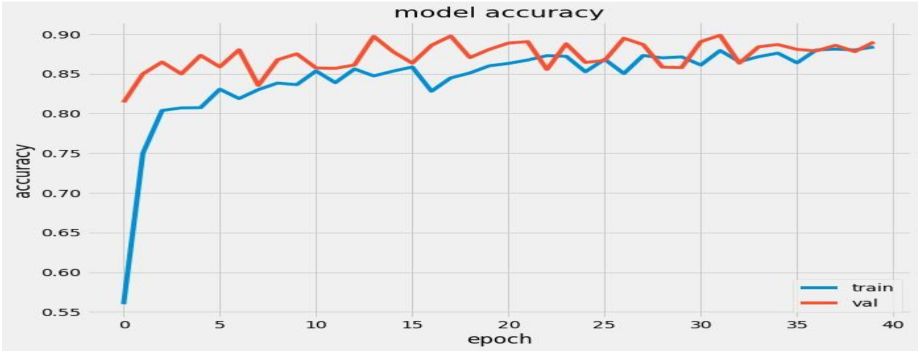


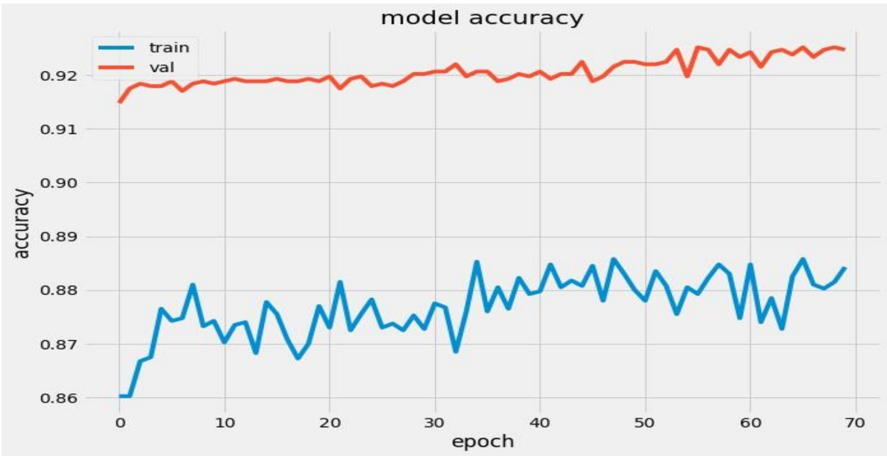
2. Model After Tunning

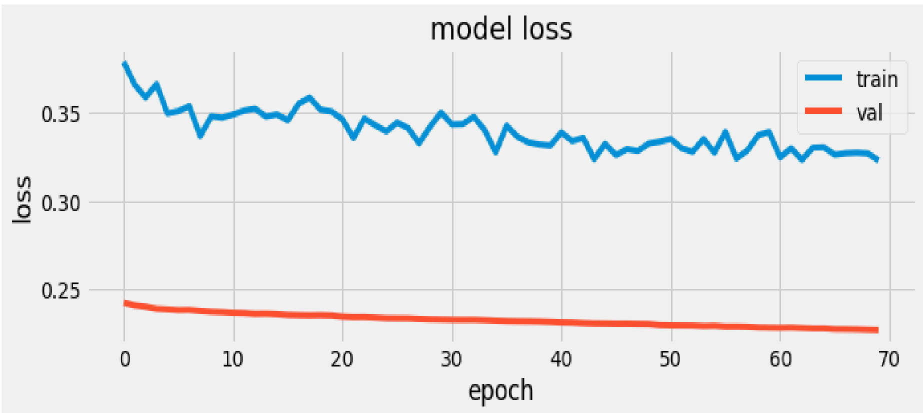
Accurate tuning means taking the weights of a trained neural network and using it as a form of a new model being trained on data from the same field (often e., pictures. It is used for:

i. Speed up training
ii. Overcoming small data sheet size


epoch vs loss
epoch vs accuracy


model accuracy
accuracy
epoch
train
val


model accuracy
accuracy
epoch
train
val


model loss
loss
epoch
train
val

4. CONCLUSION

This paper proposed a method of food image classification based on inception-v3 transfer learning in Food images, and the method was compared with other methods. In particular, the method of food image classification based on migration learning can achieve higher accuracy. Moreover, the neural network model based on transfer learning performs better in food image classification on Food 101 database than the model based on original DCNN. Fine-tuning Inception-v3 transfer learning can effectively improve the accuracy of food image classification, so the model can provide an effective and rigorous computer-assisted diagnostic when food image data is insufficiency. If the network selection of transfer learning is inappropriate, the problem of negative transfer may occur, which will lead to the decrease of accuracy and the increase of training time. Therefore, how to select the appropriate network better for food image tasks is the further research direction.

For the problem of the large gap between the specificity
and sensitivity of deep learning model in this study, we found that ensemble learning has a good performance in medical images from many pieces of literature, their sensitivity and specificity are both high. Therefore, we will make further research on ensemble learning in future research work.

REFERENCES

[1] L. Shao, F. Zhu and X. Li, "Transfer Learning for Visual Categorization: A Survey," in IEEE Transactions on Neural Networks and Learning Systems, vol. 26, no. 5, pp. 1019-1034, May 2015, doi: 10.1109/TNNLS.2014.2330900.

[2] C. Szegedy et al., "Going deeper with convolutions," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, 2015, pp. 1-9, doi: 10.1109/CVPR.2015.7298594.

[3] Neyshabur, B., Sedghi, H., & Zhang, C. (2020). What is being transferred in transfer learning? ArXiv, abs/2008.11687.

[4] Tan C., Sun F., Kong T., Zhang W., Yang C., Liu C. (2018) 'A Survey on Deep Transfer Learning'. In: Kůrková V., Manolopoulos Y., Hammer B., Iliadis L., Maglogiannis I. (eds) Artificial Neural Networks and Machine Learning – ICANN 2018. ICANN 2018. Lecture Notes in ComputerScience, vol 11141.Springer, Cham. https://doi.org/10.1007/978-3-030-01424-7_27

[5] Cheng Wang, Lin Hao ; Xuebo Liu ; Yu Zeng ; Jianwei Chen ; Guokai Zhang et. al. C. Wang et al., "Pulmonary Image Classification Based on Inception-v3 Transfer Learning Model," in IEEE Access, vol. 7, pp. 146533-146541, 2019, doi: 10.1109/ACCESS.2019.2946000.

[6] S.Y. Ge, Z.H. Zhu, Hong Liu. "Comprehensive evaluation method for distribution network fault risk". In the Journal Of Proceedings of the CSU-EPSA 26(07) (2014).

[7] B. Wen, Z.Y. Liu, B.Z. Hong. "Discussion on impact of weather on reliability of distribution networks". In the Journal Of Guangdong Electric Power 25(09) (2012).

[8] W. Zhang, W.X. Sheng, K.Y. Liu. "A prediction method of fault risk level for distribution network considering correlation of weather factors". In the Journal Of Power System Technology, 42(08) (2008).

[9] H. Liu, R.A. Davidson, A.M. Asce. "Negative Binomial Regression of Electric Power Outages in Hurricanes". In the Journal Of Journal of Infrastructure Systems, (2005).

[10] D.T. Radmer, P.A. Kuntz, R.D. Christie. "Predicting Vegetation Related Failure Rates for Overhead Distribution Feeders". In the Journal Of IEEE Power Engineering Review, 22(9) (2007).

[11] S.W. Fei, Yu Sun. "Fault prediction of power transformer by Combination of Rough Sets and Grey Theory". In the Journal Of Proceedings of the CSEE, 28(016) (2008).

[12] A. Pahwa. "Modeling Weather-Related Failures of Overhead Distribution Lines". In proceedings of Power Engineering Society General Meeting, (2017).

[13] Dian He, D.D. Tan, H.L. He. "Quantity prediction of distribution network faults based on Big Meteorological Data". In the Journal Of Journal of Jishou University, 040(001) (2014).

[14] Wei Xi, Peng Li, X.B. Guo. "Application of Correlation Analysis Method based on multidimensional time series in prediction of equipment failure". In the Journal Of Advances of Power System & Hydroelectric Engineering, 000(012) (2014).

[15] L. Breiman. "Random Forests". In the Journal Of Machine Learning, 45(1) (2001).

[16] A.P. Sanil. "Principles of Data Mining". In the Journal Of Publications of the American Statistical Association, 98(461) (2007).

[17] J.W. Han, M. Kamber, J. Pei. "Data mining: Concepts and techniques". In the Journal Of The Morgan Kaufmann Series in Data Management Systems, (2011).