Adaptive Learning Rate based Convolutional Neural Network for Food Recognition

Urvashi Rakholiya[1], Chirag patel[2], Ankita Gandhi[3], Hetal Bhaidasna[4]

Department of Computer Engineering

Parul University of Engineering and Technology (PIET), Vadodara

Email id - *patelurvashi164@gmail.com[1],chirag.patel8781@paruluniversity.ac.in[2], ankita.gandhi@paruluniversity.ac.in[3],hetal.bhaidasna@paruluniversity.ac.in[4]*

DOI:- https://doi.org/10.47531/SC.2022.29

Abstract

Picture recovery and characterisation in the food field have increasingly concerned examination points in interactive media investigation and applications. In ongoing years, with the quick advancement of the Internet business and sight and sound innovation, picture order and recovery innovation has become an exploration hotspot at home and abroad. Based on this, this paper proposes research on food image classification and image retrieval methods based on visual features, machine learning, CNN, R-CNN network and uses an adaptive learning rate is presented for training neural networks. Most probably conventional updating methods in which the learning rate gradually decreases and increases during the training period. Experiments with well-known datasets to train a multilayer perceptron show that the proposed method effectively obtains better test accuracy under certain conditions.

Keywords:- *Deep Learning, CNN; Food Dataset; Food Recognition; Machine Learning; CNN Classifications.*

INTRODUCTION

A. Image classification

Image classification is the most important part of computer vision. In simple words, classification is the categorisation of the same kind of data in the same category. Image classification is a process that includes image pre-processing, image segmentation, key feature extraction [1]. There are three methods of classification supervised learning, unsupervised learning, semi-supervised learning.

B. Deep neural network

A deep neural network (DNN) is an artificial neural network with multiple layers between the input and output layers. A deep neural network is also called "stacked neural networks"; networks composed of several layers. A deep neural network performs automatic feature extraction without human intervention. The deep learning algorithms are:

1. Convolutional Neural Network
2. Deep Boltzmann Machine
3. Deep Belief Network

4. Deep stacking Network
5. Recurrent Neural Network

C. Convolutional neural network

A convolutional neural network (CNN) is a specific type of deep neural network which is especially useful for image classification and image recognition. A convolutional neural network is also known as a convNet. CNN can handle a large amount of data and estimate the features automatically and utilise it for image classification [2]. A convolutional neural network has input, output, and hidden layers like other artificial neural networks. Some of these layers are the convolutional layer, pooling layer and fully connected layer. In CNN, the layers are organised in three dimensions: Width, height and depth. The main importance of CNN is that the neurons in one layer do not connect to all the neurons in the next layer but only to a small region. Last, the final output will be reduced to a single vector of probability scores, organised along the depth dimension. The CNN has an excellent performance in machine learning problems [3]. Especially the applications that deal with image data, computer vision. See figure 1

Steps for build a Convolutional Neural Network:

i. Convolution layer

This is the first layer of a convolutional neural network. In this layer, perform a mathematical operation that takes inputs as image matrix and filter or kernel. The output of a convolution layer is called a feature map.

ii. Pooling (Subsampling) layer

The pooling layer reduces the size of the feature map, which is created by the previous layer.

iii.fully- connectedlayer

This layer is the same as the regular neural network layer, which takes input from the previous layer, computes the class scores, and outputs the 1-D array equal to the number of classes.

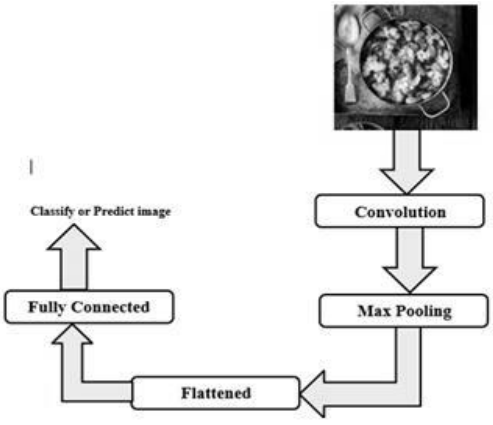


Fig 2: Example of CNNs process

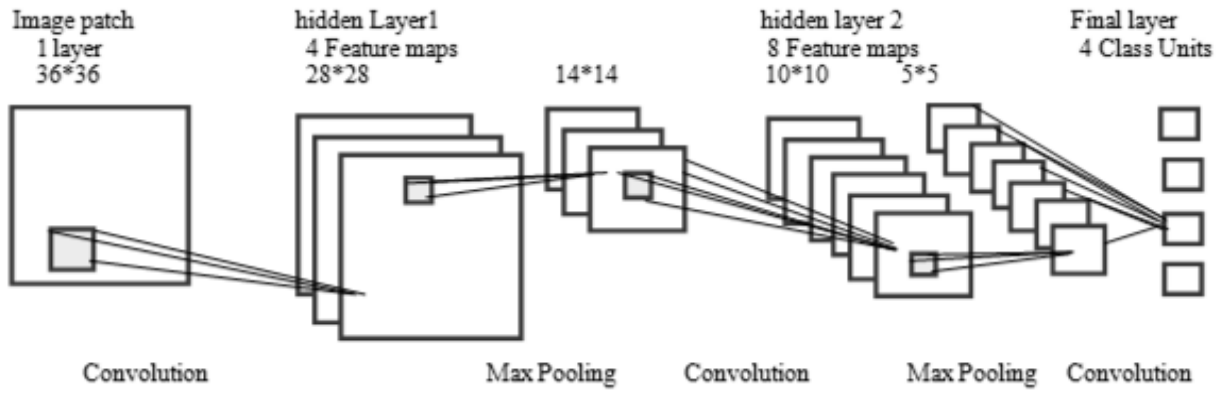


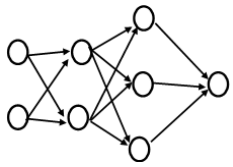
Fig 1: Convolution Neural Network Process

D. Recurrent neural network(RNN)

A recurrent neural network (RNN) is one type of deep neural network for sequentialdata. This is the first algorithm that remembers its input due to an internal memory, ideally suited for machine learning problems involving sequential data [4].

A recurrent neural network cannot remember from longer sequences or time. In an RNN, the information cycles through a loop. When it makes a decision, it considers the current input and wh at it has learned from the previous inputs [4].

Recurrent Network



Input Layer Hidden LayerOutput Layer

Fig 3: Recurrent Nural Network

Nowadays, being fit has become an important life goal of people and to help them out in this journey, various e-health applications have been developed. These applications use image processing methods to detect food present in a particular cuisine. Therefore, to overcome this, various algorithms were implemented for food image recognition to food items intake by a particular user. The analysts used algorithms like Convolutional Neural Network (CNN), RNN etc. They came up with an easy and intelligent solution for measuring food intake and collecting their dietary information. To recognise multiple food items in one picture and calculate their learning rate requires high skills and sound knowledge of the pros and cons of every method tried by different researchers. It is indeed a difficult task as it includes a collection of huge datasets for training the system. It was also observed that there were significant intraclass variations in the captured images of the food items [2].

This paper consists of the comparisons between various algorithms to understand the merits and demerits of each and develop an application providing higher accuracy in results.

LITERATURE SURVEY

For recognising multiple food items in the images captured using mobile devices Pengcheng Wei, Bo Wang, [1] based on Faster R-CNN network, implemented on food data set Dish-233 and 48, 189 images. Other compared with CNN-GF, the performance is improved by 5%. This method is more discriminative visual features and improved in food image retrieval and classification tasks.

AfsanaAhsanJeny, Masum ShahJunayed, Ikhtiar Ahmed, Md. Tarek Habib [2] used different classification algorithms to predict data from the Food dishes' dataset. (fig. 4) These algorithms are Adam optimisation algorithm Confusion Matrix performed on Deep Residual Neural Network for grouping six classes of food also used two convolutional neural networks separately for grouping six classes of food images and achieved the highest accuracy of 98%.

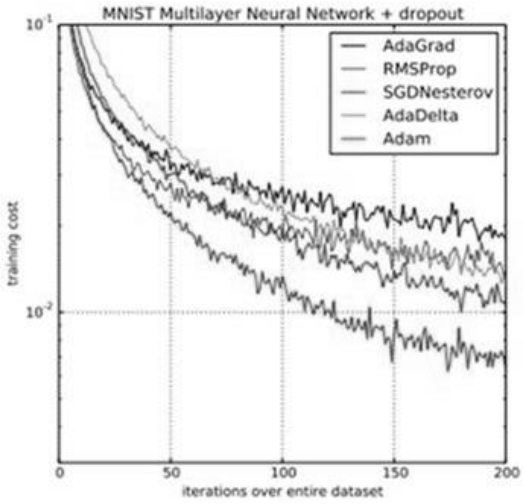


Fig 4: Adam optimazation Grap

V. Hemalatha Reddy, Soumya Kumari, Vinitha Muralidharan, Karan Gigoo, and Bhushan S. Thakare[3] proposed classification methods like Stochastic gradient and region selection, K-mean clustering, and segmentation based on Graph Cut etc. These methods are used to detect the food items, calculate calories measurements, and analyse the food data set. Gradient: The stochastic gradient is generally used to minimise the error of a model on the training data. The neural network is trained to find the weight (w) and bias (b) to minimise the quadratic cost function. This is done by defining the cost function as

$$\text{Cost}(w, b) \equiv 1/2 \parallel y(x) - a \parallel \quad (1)$$

where, w → collection of all the weights in the network, b → biases, a → vector of outputs from the network, x → input. Using this function improves the cost function by making it easier to make minor changes in the weights (w) and biases (b).

Deep learning is a growing method that helps a computer model perform classification tasks directly from images. It's a machine learning approach and has been executed to discover multiple levels of representation, including (Fig. 1).

TomoumiTakase, Satoshi Oyama, MasahitoKurihara.[4] proposed an approach to find a prediction of food dishes and find the learning rate using Stochastic gradient descent Algorithm, ALR technique (Adaptable Learning Rate Tree algorithm). In the proposed model, the author compares the result of increases or decreases the learning rate adaptively so that the training loss decreases as much as possible. The comparison shows the ALR algorithm can obtain greater efficiency and accuracy.

Weishan Zhang Dehai Zhao Wenjuan Gong Zhongwei Li Qinghua Lu Su Yang [5] used the Support Vector Machine method. Implemented on food data set, the UEC-Food101 dataset, is an open 100-class food image dataset including about 15000 images, and the other is a fruit dataset established by ourselves including over 40000 images. We have achieved the best accuracy of 80.8% on the fruit dataset and 60.9% on the multi-food dataset.

Yi Sen Ng, WanqiXue, Wei Wang, Panpan Qi [6] used different classification algorithms Like Stochastic Gradient Descent (SGD), Confusion matrix, Augmentation technique to predict data from Food da. These algorithms are Adam optimisation algorithm Confusion Matrix performed on Deep Residual Neural Network for grouping six classes of food also used two convolutional neural networks separately for grouping six classes of food images an achieved the highest accuracy of 98%.

Marc Bola˜nos, Marc Valdivia, and PetiaRadeva, [7] proposed classification methods like CNN and RNN (Convolutional Neural Network and Recurrent Neural Network). These methods are used to detect the food items applying image food recognition by using the restaurants' menu and improving the baseline by 15%.

SandhyaArora, Gauri Chaware, Devangi Chinchankar, Eesha Dixit and Shevi Jain [8] used SIFT and HOG. BOF (Bag-of-Features) and DCNN. In the proposed model, the author compares increases or decreases accuracy and use of hand-crafted features s. Food items are being recognised using cutting sounds, acoustic sensors, so on.

Lei Zhou, Chu Zhang, Fei Liu, Zhengjun Qiu, and Yong He [9] used Faster R-CNN and Image retrieval method based on BOW (Bag of Words) classified SIFT features and implemented on food data set Food101 the combination of deep learning and multisource data fusion including RGB images, spectra, smell, taste, and considered to make a more comprehensive assessment of food, analysed obtained the best classifier.

Abdulkadir ŞENGÜR, Yaman AKBULUT, Ümit BUDAK, [10] used different classification algorithms like Support Vector Machine (SVM) and fine-tuning a pre-trained CNN. These methods used to feature size 4096 are extracted from fc6 and fc7 layers and concatenated with various combinations to determine the best deep feature sequence for food image classification. Achieved the accuracy of 79.85%

Mohammed A. Subhi, Sawal Md. Ali [11] proposed classification methods like CNN (Convolution Neural Network). These methods are used to detected the food items on the food dataset and calculated the accuracy.

Niki Martinel, Gian Luca Foresti, Christian. [12] proposed an approach to using computer vision and pattern recognition algorithms and find a prediction of food dataset-100. Archived an accuracy of 90.27%.

PROPOSED WORK

The methodology outlined the details of tools and techniques used for collecting food image data, methods used to analyse food image classification and followed for the present research study. The methodology consists of a selection of food image classification and proposed work are as shown in figure find the learning rate with the help of the Activation function (fig. 5)

The steps undertaken to apply CNN algorithms shown in fig 2 are presented as follows-

a) Activation Function:

1. Sigmoid Function

It is also known as the logistic activation function. The sigmoid function is mainly used for the output layer of models, which predicts the probability as an output. The range of sigmoid is from 0 to 1.

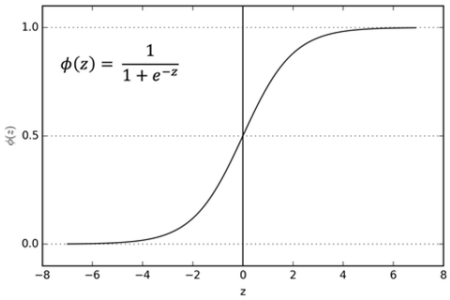


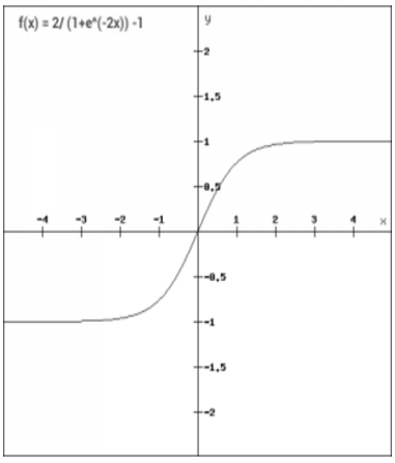
Figure 5: Sigmoid function

Equation of sigmoid function

F(x) =1/(1+e^-x) (2)

2. Tanh Activation Function

Tanh Function does the negative inputs will be mapped strongly negative, and zero inputs will be mapped near zero. The range of Tanh is from -1 to1.



Equation of Tanh function

$F(x) = (2/(1+e^{-2x}))-1$ (3)

See figure 7

One of the most important issues that occur during the detection of images is understanding the various textures that differentiate the different parts of the image. These segments and textural differences can be found out using the filter.

CONCLUSION

This paper has discussed various techniques proposed by different authors to detect and recognise multiple food items placed on a plate. We saw how each method work, how much accuracy these methods give and concluded that Convolutional Neural Network shows greater accuracy than any other base methods for food detection, we can conclude that through the proper set of food image, classification and using various Convolution neural network techniques, accuracy and Learning rate can be obtained efficiently. It will give better results compared to another classifier method.

In our future work, we will develop an application to detect and recognise food items present in that dish, analyse the papers, and give valuable results. It decreases/ increases the learning rate compared to other techniques.

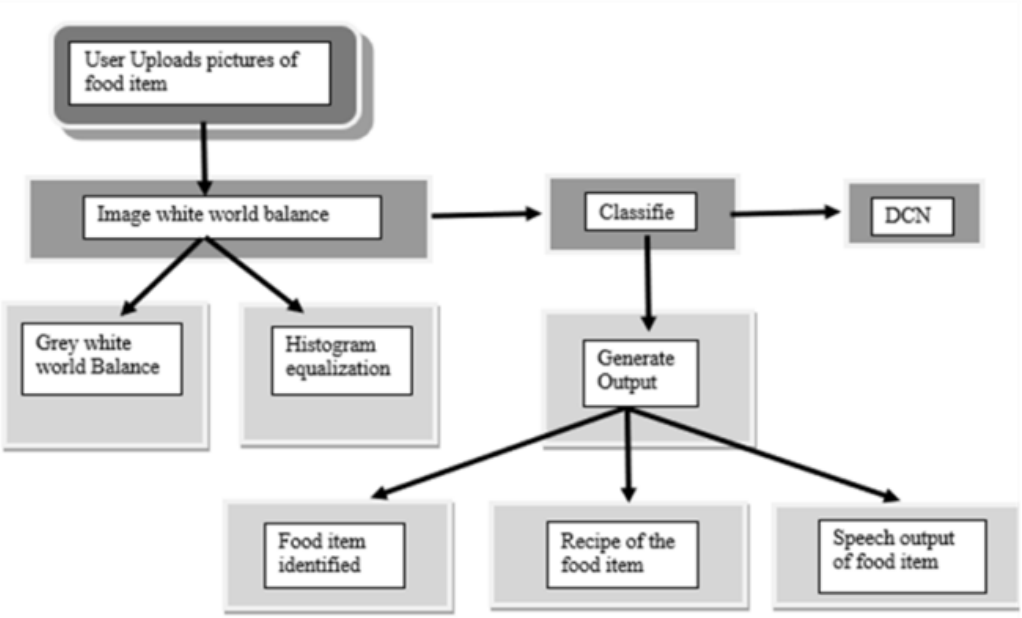


Fig 7: Proposed work

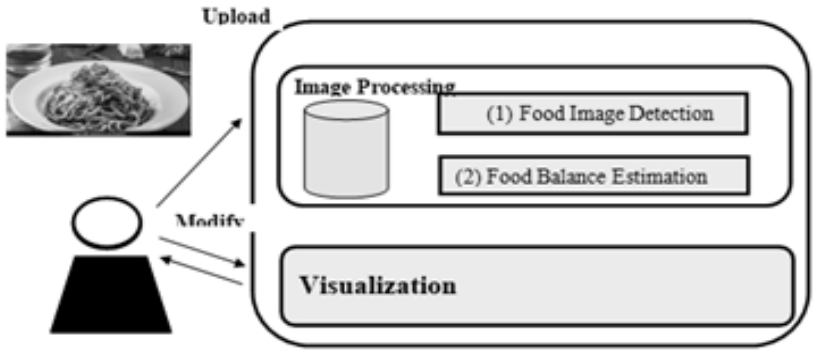


Figure 8: Functional overview of food

R.No, Year	Algorithm/classifier/ method used	Accuracy	Disadvantage/ limitation	Future scope	Data set used
[1] (2020-IEEE)	- Faster R-CNN, - Image retrieval method based on BOW (Bag of Words) classified SIFT features.	95%	- The biggest disadvantage of this algorithm is the dimensional disaster.	- Experiments on more large-scale food data sets verify the scalability of the method in this paper. - Consider more contextual information of food images, such as geographic location information, etc.	FoodDataSet
[2] (2019-IEEE)	- Adam optimisation algorithm. - Confusion Matrix	98.16%	The main drawback is inception v3 always better performed on three main food image datasets than other datasets.	Remove this little overfitting. .	UEC FOOD 100
[3] (2019-IEEE)	- Stochastic gradient and region selection. - K-mean clustering and segmentation based on Graph Cut	90%	- Too much variance in food appearance Multiple food detection in composite meals – Users acceptance studies. - Component costs (thermal cameras) are highly expensive.	Develop an application to detect and recognise food and calculate total calorie counts present in that dish.	FOODD Dataset
[4] (2018-IEEE)	- Stochastic gradient descent Algorithm - ALR technique (Adaptable Learning Rate Tree algorithm	91%	These methods are that they have sensitive hyperparameters which are difficult to tune appropriately. - Learning rate is	Developing its efficient extension for individually controlling the learning rate for each weight as in AdaGrad is also a possible future work.	FOOD101

			automatically controlled by reinforcement learning.		
[5] (2016-IEEE)	Support Vector Machine	80.8%	- RGB multi-food images have not gotten an ideal result. - CNN does not work well for a small-scale dataset.	A large-scale dataset is needed, with an accuracy of 100%, which will enlarge the recognition error rate. Therefore, we can build a larger dataset by manual selection to get a better result.	UEC-FOOD 101
[6] (2019 - IEEE)	- Stochastic Gradient Descent (SGD) - Confusion matrix - Augmentation technique	98%	Low latency	Image augmentation techniques that do not alter shapes such as zoom, flip, and width height shift are more beneficial than other augmentation techniques that change shapes; various augmentation techniques can also be combined for better results.	FOOD101
[7] (2019 - IEEE)	CNN and RNN (Convolutional Neural Network and Recurrent Neural Network)	75%	The algorithm is dealing with a wide variety of names. - Dishes with exotic names, which our language model cannot easily learn.	- Introduce the GPS information of the images. The user's location gives us a list of two or three candidate restaurants where they are eating.	FOOD Dataset
[8] (2019, IEEE)	- SIFT and HOG. - BOF (Bag-of-Features)	--------	Could be allowed to add new training images and food categories	try and improve the accuracy rates.	FOOD101

	- DCNN				
[9] (2019, IEEE)	- CNN (WISeR) and RGB. - DNN	85%	--	--	FOOD101
[10] (2018, IEEE)	- Support Vector Machine (SVM). - Fine-tuning of a pre-trained CNN.	85%	CNN does not work well for a small-scale dataset.	Planning to investigate the effect of the early layers feature concatenation on food category classification.	----
[11] (2018, IEEE)	CNN (Convolution Neural Network)	---	----	Increase the accuracy of model classification a deep network use to multi-dimensional model.	---
[12] (2016, IEEE)	computer vision and pattern recognition algorithms.	90.27%	substantial memory loads as well as significant computational efforts to process a single deployment.	Compress the network to obtain a DCNN architecture or to exploit binary weights.	FOOD dataset-100

REFERENCES

1. Pengcheng Wei, Bo Wang, 'Food image classification and image retrieval based on visual features and machine learning,' Computer Science and Technology-2020, IEEE.
2. Afsana Ahsan Jeny, Masum ShahJunayed, Ikhtiar Ahmed, Md. Tarek Habib 'FoNet-Local Food Recognition Using Deep Residual Neural Networks', International Conference on Emerging Research in Electronics, Computer Science and Technology-2019, IEEE.
3. V. Hemalatha Reddy, Soumya Kumari, Vinitha Muralidharan, Karan Gigoo and Bhushan S. Thakare, 'Literature Survey—Food Recognition and Calorie Measurement Using Image Processing and Machine Learning Techniques, International Conference on Emerging Research in Electronics, Computer Science and Technology-2020, IEEE.
4. Tomomi Takase, Satoshi Oyama, Masahi to Kurihara, 'Effective neural network training with adaptive learning rate based on training loss', International Conference on Emerging Research in Electronics, Computer Science and Technology-2018, IEEE.
5. Weishan Zhang Dehai Zhao Wenjuan Gong Zhongwei Li Qinghua Lu Su Yang,' Food Image Recognition with Convolutional Neural Networks, International Conference on Emerging Research in Electronics, Computer Science and Technology-2016, IEEE.
6. Yi Sen Ng, Wanqi Xue, Wei Wang, Panpan Qi,' Convolutional Neural Networks for Food Image Recognition: An Experimental Study, International

Conference on Emerging Research in Electronics, Computer Science and Technology-2019, IEEE.

7. Marc Bola˜nos, Marc Valdivia, and Petia Radeva,' Where and What Am I Eating? Image-Based Food Menu Recognition', International Conference on Emerging Research in Electronics, Computer Science and Technology-2015, IEEE.
8. Sandhya Arora, Gauri Chaware, Devangi Chinchankar, Eesha Dixit and Shevi Jain,' Survey of Different Approaches Used for Food Recognition, International Conference on Emerging Research in Electronics, Computer Science and Technology-2019, IEEE.
9. Lei Zhou, Chu Zhang, Fei Liu, Zhengjun Qiu, and Yong He, 'Application of Deep Learning in Food: A Review', International Conference on Emerging Research in Electronics, Computer Science and Technology-2019, IEEE.
10. Abdulkadir ŞENGÜR, Yaman AKBULUT, Ümit BUDAK,' Food Image Classification with Deep Features', International Conference on Emerging Research in Electronics, Computer Science and Technology-2018, IEEE.
11. Mohammed A. Subhi, Sawal Md. Ali, 'A Deep Convolutional Neural Network for Food Detection and Recognition', International Conference on Emerging Research in Electronics, Computer Science and Technology-2018, IEEE.
12. Niki Martinel, Gian Luca Foresti, Christian Micheloni'Wide-Slice Residual Networks for Food Recognition, International Conference on Emerging Research in Electronics, Computer Science and Technology-2016