

# Continual Learning for Food Recognition Using Class Incremental Extreme and Online Clustering Method: Self-Organizing Incremental Neural Network

Rashida Bansiwala<sup>1</sup>, Prof. Pramod Gosavi<sup>2</sup>, Prof. Rahul Gaikwad<sup>3</sup>

<sup>1</sup>PG Student, Godavari College of Engineering, Jalgaon, Maharashtra, India

<sup>2,3</sup>Assistant Professor, Godavari College of Engineering, Jalgaon, Maharashtra, India

*Received on:* 16 July, 2021

*Revised on:* 11 August, 2021

*Published on:* 13 August, 2021

**Abstract-** Nowadays, standard intake of healthy food is necessary for keeping a balanced diet to avoid obesity in the human body. Literature has indicated that accurate dietary assessment is very important for assessing the effectiveness of weight loss interventions. However, most of the existing dietary assessment methods rely on memory. With the help of pervasive mobile devices and rich cloud services, it is now possible to develop new computer-aided food recognition system for accurate dietary assessment. But Food Recognition does not allow data incremental learning and often suffer from catastrophic interference problems during the class incremental learning. This is an important issue in food recognition since real-world food datasets are open-ended and dynamic, involving a continuous increase in food samples and food classes. Model retraining is often carried out to cope with the dynamic nature of the data, but this demands high-end computational resources and significant time. This paper proposes a new open-ended continual learning framework by employing transfer learning on deep models for feature extraction, Relief F for feature selection, and a novel adaptive reduced class incremental kernel extreme learning machine (ARCIKELM) for classification. Relief F reduces computational complexity by ranking and selecting the extracted features. The novel ARCIKELM classifier dynamically adjusts network architecture to reduce catastrophic forgetting. It addresses domain adaptation problems when new samples of the existing class arrive. Results show that the proposed framework learns new

classes incrementally with less catastrophic inference and adapts domain changes while having competitive classification performance.

**Keywords-** Food recognition, ReliefF, continual learning, ARCIKELM

## I - INTRODUCTION

In the open-ended continual learning, the new images of existing classes arrive continuously, and novel classes always appear. Two types of incremental learning are an important component in open-ended continual learning:

- 1) Data incremental learning
- 2) Class incremental learning.

The data incremental learning improves the recognition performance of existing classes and adapt domain changes using newly available images. In contrast, class incremental learning continuously gains knowledge from novel classes. The importance of open-ended continual learning in many real-world recognition problems, including food recognition as the dataset is dynamic, and new concepts of interest occur over time [1]

## II- PROBLEM DEFINITION

Deep learning networks for food image classification have achieved state-of-the-art performance on various food datasets. Despite this, these models have assumed fixed datasets which have increased the gap between laboratory and real-world environment. Most of the image datasets in real life are open ended and dynamic. However, a well-trained deep learning model tends to forget the previous information while learning new information this concept known as catastrophic forgetting. This means a model learned on fixed food datasets cannot be trained easily on additional samples of current food classes or new food classes without significantly affecting its previous performance. These two problems hinder the usage of deep learning models for open-ended continual learning. To understand the catastrophic forgetting during incremental learning, two theories have been proposed. The first theory suggests that human plasticity decreases on neurons that have learned previous information. This decreased plasticity on neurons helps to retain previous information. The second theory explains that human extracts high-level information and stores that information in different brain areas while retaining episodic memories.

### III - RELATED WORK

Most of the existing approaches for food recognition assume that the training dataset has all the classes and variations within them. However, in a real-world setting, the limited number of classes and images of existing food classes are initially available. Therefore, the food recognition system must satisfy the criteria of open-ended continual learning. It should be able to learn new classes incrementally in real-time and does not need to store exemplary images datasets this will increase the complexity of the model and data security concerns. The online learning should be considered to adapt variation in the existing class images. This section briefly discussed the present literature on food recognition in context with mentioned challenges. It includes a wide range of methods, from simple Euclidean distances to spatial pyramid convolution neural networks for food recognition.

Below are the different methods:

1. A. Distance Based Method
2. Probabilistic Model And Decision Tree
3. Support Vector Machine
4. Convolutional Neural Network
5. Extreme Learning Machine

### IV- THE APPROACH

#### A. System Overview

The proposed framework takes into account two attributes of open-ended learning 1) Class incremental learning 2) Data incremental learning. It consists of three modules A. Feature extraction module B. Feature selection and C. Classification. Fig. 1 show the flow chart of the proposed architecture. During a training period, each incoming image is given to the selected deep feature extractor module that extracts the features. They are then ranked using the Relief F method, and the best features are chosen based on our proposed strategy. The selected features from the Relief F method are considered as the final representation of the image. Although, the extracted features from the deep model have good generalization ability. The fixed class architecture softmax for final classification does not take advantage of this ability. This study uses novel ARCIKELM which learns from these representations while satisfying both criteria of open-ended continual learning. It adds new output and hidden neurons when image representations belong to the novel class. In case they are from existing classes, it updates the model sequentially by our proposed strategy and only adds new hidden neurons when required.

During the classification stage, the features are extracted from the test image by using the same deep feature extractor. The best representations are chosen based on the ranking of the features from the Relief F method. Finally, ARCIKELM makes the final decision.

#### B. Image Feature Representation

Deep learning models for food recognition can learn the best image representations and eliminate the need for handcraft feature extraction process which is based on prior knowledge. Our proposed framework applies transfer learning with online data augmentation on pre-trained image-net deep model which is then used to extract features. In transfer learning a model created for one task is used again as the base point for a second task. When modeling the second task, it enables for fast advancement and enhanced results. It is prominent in deep learning for feature extraction because the training of the model from scratch requires specialized computational resources such as GPU and long training time. This is the major hindrance for real-time incremental learning. To determine the best architecture of a model for food feature extraction, I have explored three state-of-the-art deep learning networks, ResNet-50 [9], DenseNet-201 [10], and Inception ResNet-V2 [11]. The Image Net weights initialize the models. They then are \_ne-tuned using food datasets and the resultant model is used for feature extraction. The experiment section

has discussed the results.

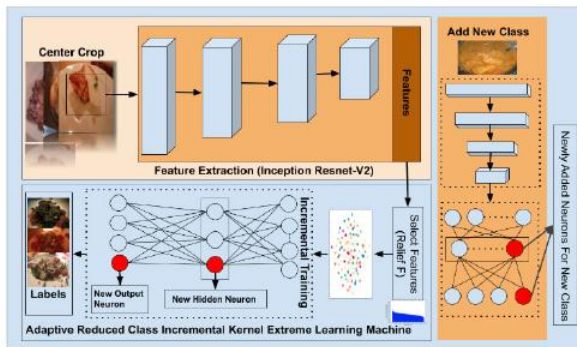


Fig 1- Flow chart of proposed framework. (The red circle shows neurons which are incrementally added to the network.)[2]

**Data Augmentation:** Data augmentation generates transformed versions of pictures that are in the same class as the initial picture in the training. Transformations include a variety of image processing activities like zoom, horizontal shift, rotation, etc. The objective is to add fresh, credible instances to the training. This implies, creating different variants of training that the model is highly likely to detect. For example, the horizontal flip of food is a plausible scenario because the picture could be taken in either direction. However, the vertical flip of food would produce a useless result as a picture of food turned upside down does not make any sense. The study has applied 'shifting', flip, zoom range, Channel Shift Range and fill mode. They are explained in the following paragraph. In shifting, image pixels are moved in the horizontal or vertical direction. The group of pixels which are deleted in one region is copied into another region of the image. This study uses both horizontal and vertical shifts. Flipping includes inverting the rows or columns of pixels. The vertical flip reverses the columns of pixels and the horizontal flip reverses the rows of the pixel. The experiments have investigated the horizontal flip as a vertically flipped image of food does not make any sense. In zoom augmentation, the picture is zoomed in or out, either by adding fresh pixel values around the picture or by interpolating pixel values respectively. It is important for food images as various users take images at different zoom levels. Channel shifting is the method of capturing the pixel's red, green or blue values in a picture and adding those values to pixels in distinct positions in the picture. The experiments have shifted channel by 30 degrees using online data augmentation. Finally, in the fill augmentation, fresh pixels are introduced in the picture that is not included in the initial picture boundaries. For example, after rotation black area in the corners would be added in the picture that was not included in the original picture. The Fill mode

specifies how I manage these regions. This is handled by either using a constant values method or default method. The constant value is applied to all the new pixels and sub-pixels. In the method, all the new pixels are either given the black value or white. To fill the area with black, 0 is set as the default value and to fill the area with a white, default value is set as 1. Fig.2 illustrates the different modes of online data augmentation applied in this study on a randomly selected food image.

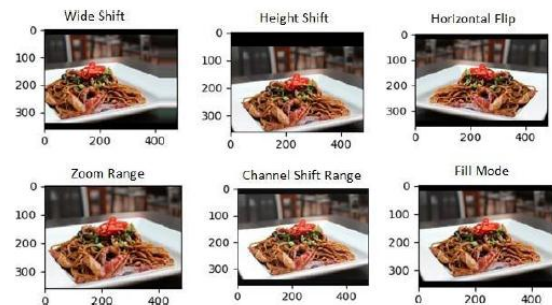


Fig: 2 Different mode of data augmentation

The results in the experimental section discuss in detail the impact of data augmentation. Based on the results in the experimental section, the pro-posed framework has used Inception-Resnet-V2 for feature extraction after applying transfer learning with online data augmentation. It is the first step in the food recognition process. The next section explains Feature selection.

### C. Image Feature Selection

The extracted features from the deep learning model have a very high dimension and there are only subsets of features that are relevant in determining the results. These low ranked features increase the computational complexity of classifiers. This study uses Relief F to rank features and selects the best features. It is the only algorithm for detecting feature dependencies that do not search through feature combinations but uses the concept of the nearest neighbor to detect feature dependencies. This study has evaluated three configurations for selection of features and determined which configuration has higher accuracy and less training time for food. Finally, for classification purposes, it has used the best configuration. The experimental section presents the results of the feature selection.

### D. Classification Methods

The final step in our proposed model is incremental learning and recognition. Most of the existing studies use batch learning approaches for the classification of food, which requires fixed classes, making them unsuitable for

real-time settings. This paper evolves the knowledge in the area of extreme learning machines for continuous recognition of food dishes. For this purpose, this study has proposed a novel method that learns incrementally with less training time, higher accuracy, and better generalization ability. This section has discussed the basics of class incremental extreme learning machine, followed by our proposed method. The notations for symbols and acronym definition in this paper are summarized in Table 1 and Table 2.

**TABLE 1.** Notation definition.

Symbol	Description
$X_b; X_n$	Observed samples of based and new classes
$k_b; k_n$	Initial kernel and updated kernel matrix
$\beta^{(b)} \beta^{(n)}$	Initial output weights and updated output weights
$M$	Transformation matrix
$N$	Neuron eligibility matrix
$\mu_{th}$	Membership value

**TABLE 2.** Acronyms definition.

Symbol	Description
ELM	Extreme learning machine
KELM	Kernel extreme learning machine
CIELM	Class incremental extreme learning machine
ACIELM	Adaptive class incremental extreme learning machine
ARCIKELM	Adaptive reduced class incremental kernel extreme learning machine

#### IV- PROPOSED WORK

The online clustering method like self-organizing incremental neural network can be used to select the mapping nodes which best represent the class for ARCIKELM, and help to select the nearest nodes during classification. This can reduce catastrophic forgetting and is noise invariant when the input for classification is far away from existing neurons. The other direction is the auto-scaling of computational resources in a cloud environment.

#### V- CONCLUSION

The Food recognition dataset is open-ended and dynamic. Continuously increasing in food samples and classes. A deep learning model for food recognition in existing system assumes that classes exist initially for all the food classes and variations. It causes catastrophic forgetting during class-incremental learning. This research study addresses these challenges by proposing a new framework of open-ended continual learning for food recognition. We aim to reduce catastrophic forgetting by using a hybrid scheme for open-ended continual learning. The online clustering

method like self-organizing incremental neural network can be used to select the mapping nodes which best represent the class for ARCIKELM, and help to select the nearest nodes during classification. This can reduce catastrophic forgetting and is noise invariant when the input for classification is far away from existing neurons. We will use state-of-the-art deep learning networks to extract features, Relief F for feature ranking and selection, and ARCIKELM for classification. Inception-Resnet-V2 has superior performance than any other features. However, features extracted from deep learning models have a very high dimension and increases classification time. The framework has used the Relief F method to determine the optimal length and found that the Relief F method has reduced the accumulative learning time of the proposed classifier for all the datasets by 52.14%. For addressing the challenge of data incremental learning and class incremental learning, the framework used the novel Adaptive reduced class incremental kernel extreme learning machine (ACIELM). It dynamically increases hidden neurons and output neurons. The decreased plasticity of previous neurons reduces catastrophic forgetting. Experimental results on five catastrophic forgetting measures and four classification performance measures demonstrated that the proposed classifier has superior performance as compared to existing ACIELM and CIELM. As the novel classes and new images arrive, the computational resources required increases, it require auto scale its computational resources in a cloud environment. The comparison of the proposed framework with other architectures for food recognition like supervised extreme learning committee, PMTS, GTBB etc. show competitive performance while satisfying the criteria of open-ended continual learning.

#### REFERENCES

- [1] 'An Open-Ended Continual Learning for Food Recognition Using Class Incremental Extreme Learning Machines' Ghalib Ahmed Tahir And Chu Kiong Loo , (Senior Member, IEEE) , April 5, 2020 S. Rousseau, Food 'Porn' in Media. Dordrecht, The Netherlands: Springer, Nov. 2014, doi: 10.1007/978-94-007-0929-4\_395.
- [2] C. Chen, W. Min, X. Li, and S. Jiang, "Hybrid incremental learning of new data and new classes for hand-held object recognition," *J. Vis. Commun. Image Represent.*, vol. 58, pp. 138148, Jan. 2019, doi: 10.1016/j.jvcir.2018.11.009.
- [3] T. Mensink, J. Verbeek, F. Perronnin, and G. Csurka, "Distance-based image classification: Generalizing to new classes at near-zero cost," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 11, pp. 26242637, Nov.

- 2013, doi: 10.1109/TPAMI.2013.83.
- [4] D. Maltoni and V. Lomonaco, "Continuous learning in single-incremental- taskscenarios," *CoRR*, vol. abs/1806.08568, pp. 5673, Apr. 2018.
- [5] G. Ditzler, M. Roveri, C. Alippi, and R. Polikar, "Learning in no stationary environments: A survey," *IEEE Comput. Intell. Mag.*, vol. 10, no. 4, pp. 1225, Nov. 2015, doi: 10.1109/MCI.2015.2471196.
- [6] J. D. Power and B. L. Schlaggar, "Neural plasticity across the lifespan," *WileyInterdiscipl. Rev., Develop. Biol.*, vol. 6, no. 1, p. e216, Jan. 2017, doi: 10.1002/wdev.216.
- [7] G.-B. Huang, L. Chen, and C.-K. Siew, "Universal approximation using incremental constructive feed forward networks with random hidden nodes," *IEEE Trans. Neural Netw.*, vol. 17, no. 4, pp. 879892, Jul. 2006, doi: 10.1109/TNN.2006.875977.
- [8] X. Liu, C. Gao, and P. Li, "A comparative analysis of support vector machines and extreme learning machines," *Neural Netw.*, vol. 33, no. 9, pp. 5866, Sep. 2012, doi: 10.1016/j.neunet.2012.04.002.
- [9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for imagerecognition," *CoRR*, vol. abs/1512.03385, pp. 770\_778, Jun. 2015.
- [10] G. Huang, Z. Liu, and K. Q. Weinberger, "Densely connected con-volutional networks," *CoRR*, vol. abs/1608.06993, pp. 2261\_2269, Jul. 2016.
- [11] C. Szegedy, S. Ioffe, and V. Vanhoucke, "Inception-v4, inception-ResNet and the impact of residual connections on learning," *CoRR*, vol. abs/1602.07261, pp. 4278\_4284, Feb. 2016.