Incremental Learning: Deep Neural Networks

Rama Murthy Garimella, Jyothi Prasanna G C,
Vidya Sree Vankam, Maha Lakshmi Bairaju and
Manasa Jagannadan

February 2, 2020

Incremental Learning:Deep Neural Networks

Rama Murthy Garimella
Computer Science and Engineering
Mahindra Ecole Centrale
Hyderabad
rama.murthy@mechyd.ac.in

Jyothi Prasanna GC
ECE
IIIT RK-VALLEY
RGUKT AP
prasannajyothi805@gmail.com

Vidya Sree Vankam
Electronics and Communication Engineering
IIIT RK-VALLEY
RGUKT AP
vankamvidyasree@gmail.com

Maha Lakshmi Bairaju
Electronics and Communication Engineering
IIIT RK-VALLEY
RGUKT AP
lakshmiiiit49@gmail.com

Manasa Jagannadan
Electronics and Communication Engineering (ECE)
IIIT RK-VALLEY
RGUKT AP
manasajagannadan@gmail.com

***Abstract*—In this research paper, the problem of Incremental Learning is addressed. Based on the idea of extracting features incrementally using Auto-Encoders, CNNs, Deep Learning architectures are proposed. Experimental investigations are reported.**

***Index Terms*—Incremental Learning, Auto-Encoders, Convolutional Neural Networks, Classification**

I. Introduction

The research area of Computational Neuro-Science encompasses Artificial Neural Networks(ANNs). The first stage of progress on designing and implementing ANNs culminated in the back propagation algorithm utilized in the Multi Layer Percepton(MLP) implementation. The next stage of progress on ANN's was initiated with the deep learning paradigm. Specifically, Convolutional Neural Networks showed excellent progress in achieving better than human accuracy in many real world classification problems. But, CNNs are far away from being able to achieve functions performed by the human brain. For instance, the human brain acquires knowledge incrementally in classification, association and many other tasks. Thus, the human brain is endowed with "incremental learning" ability. This research paper is an effort in achieving incremental learning based on Deep Neural Networks.

This research paper is organized as follows.
In section 2, known related Literature is briefly reviewed. In section 3, CNN architectures to learn one object at a time are discussed. In section 4, auto-encoder based IL architectures are discussed. Also CNN based IL architectures are discussed.

II. Review of Related Research Literature

Human brain has the ability to acquire knowledge (classification, association, memory etc tasks) from natural physical reality INCREMENTALLY. Researchers are thus motivated to study models of INCREMENTAL LEARNING (IL) using Artificial Neural Networks(ANNs). Adaptive Resonance Theory (ART) is dedicated to enable INCREMENTAL LEARNING[3]. The first author attempted the problem using ensemble classifier models [4],[5],[6]. There were some successful results that were reported. This research paper is a culmination of such efforts. Some researchers reported Incremental Learning(IL) using Convolutional Neural Networks [1], [2].

III. Novel Convolutional Neural Network(CNN) Architecture: Learning One Object at a time :

The innovation in the CNN architecture for Incremental Learning(IL) is summarized below.

- The input images have only one object such as a CAT. It also has a dummy object so that there are 2 classes. The Convolutional & Pooling layers are trained using such input achieving good accuracy.
- A separate ANN is trained (i.e The Convolutional & Pooling layers are trained) using images containing a single different object such as a DOG and a dummy object.
- The trained architectures based on CNNs are fed to fully connected layers (Dense layers) in a parallel architecture. Such a novel architecture is fed with input images containing both a CAT and DOG separately.

The testing accuracy is determined with such ANN.

Remark: The goal is to enable the ANN to incrementally learn new objects while remembering the existing knowledge. Here we introduced different Deep Leraning architectures built with Convolutional Neural Networks and Autoencoders. These architectures perform well in the classification by providing good training and validation accuracies.

IV. Incremental Learning Architectures:

In this section, we describe various deep learning architectures we have experimented with for incremental learning.

A. *Auto-Encoder based Architecture for IL:*

Architecture-1:

In this architecture-1, we consider classification problem with finitely many classes.

Step1: Auto-Encoders (particularly convolutional) are trained (i.e. Encoder-decoder combination is utilized to extract features through nonlinear dimensionality reduction) individually for each class.

Step2: The encoder parts(for each class) are stacked in parallel and fed to fully connected layers. Such a stacked architecture is then fed with objects belonging to various classes and the validation accuracy is determined.

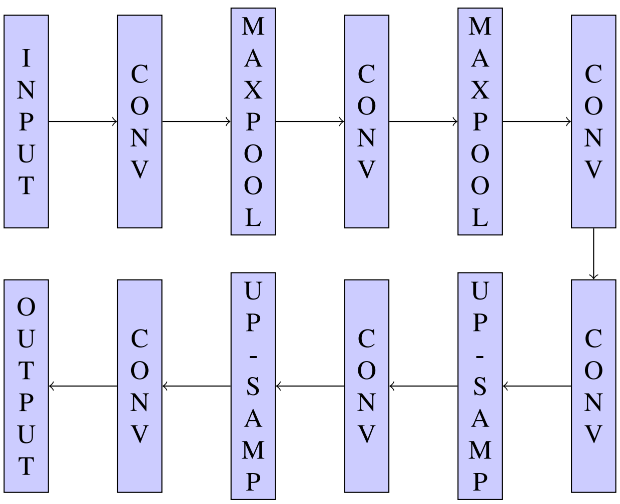

Fig. 1. Block diagram of model-1 in Architecture-1.

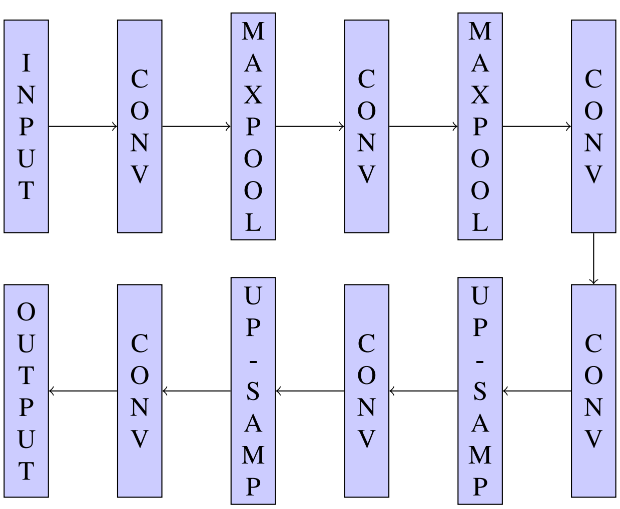

Fig. 2. Block diagram of model-2 in Architecture-1.

The architectures which we used in the model are depicted in the above diagrams. In this Convolutional Autoencoder we used 6 Convolutional layers, 3 pooling layers and 2 sampling(upsampling) layers. We used 3 x 3 kernels in the Convolutional layers and 2 x 2 kernels in the pooling(maxpool) layers. Here in this architecture of Convolutional Autoencoder we give images of shape (256,256) as input. We train the encoder and decoder part of Convolutional Autoencoder with the dataset of images belonging to one class like cats. Repeat the same procedure for the other class like dogs. Take the encoder parts from both the models and flatten them.

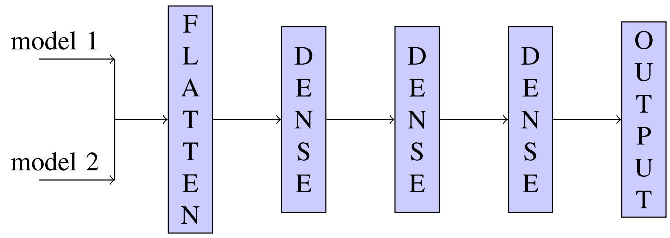

Fig. 3. Final Block diagram of Architecture-1.

Put the trained autoencoders in parallel and feed them to fully connected layers. The first fully connected layer comprises of 256 neurons with Relu as the activation function the output of which is connected to second fully connected layer with 128 neurons by using the activation as Relu. Then a dropout of 0.2 is used. The final fully connected layer is included with 2 neurons as the output with Softmax as the activation function. In the final merged model we used loss function as the binary cross entropy with optimizer Adam.

Here are the details of CNN architectures.

B. *CNN Based Architecture for IL:*

1) *Architecture-2:*

Fig. 4. Block diagram of block.

The effective idea is to train CNN's incrementally and feed them to fully connected layers for incremental classification.

The Convolution Neural Network is having five Blocks and three Dense layers including output layer. Each Block contains one Convolution layer, one Pooling layer and one Batch Normalization layer. So, this Convolution Neural Network is having five Convolution layers, five Pooling Layers, five Batch Normalization layers and three Dense layers including output layer.

In this architecture, we have used 3 x 3 as kernel size with different depth for convolution layers and Max-Pooling with 2 x 2 as kernel size. The first Dense layer is having 256 neurons, second Dense layer is having 128 neurons and Output layer contains N neurons with Softmax as activation

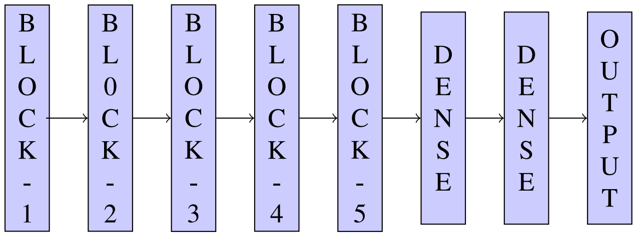

Fig. 5. Block diagram of model-1.

function (where N represents number of classes).

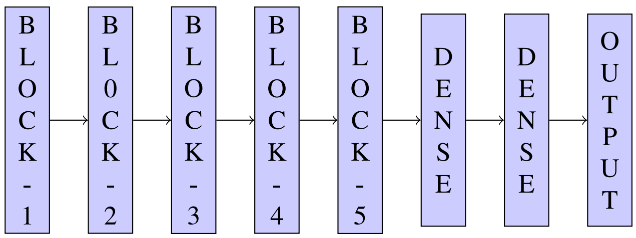

Fig. 6. Block diagram of model-2.

In this architecture-2, for two-class classification, we had taken two models . We had extracted features in CNN for a particular class separately. Then we merged these models and given their output as input to three Dense layers including Output layer. The first Dense layer is having 256 neurons, second Dense layer is having 128 neurons and Output layer contains N neurons with Softmax as activation function (where N represents number of classes).

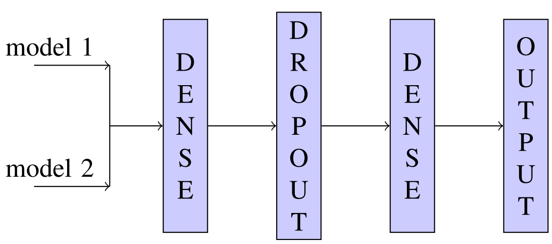

Fig. 7. Block diagram of FINAL MODEL.

2) *Architecture-3:*

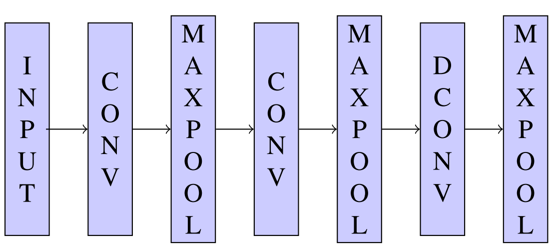

Fig. 8. Block diagram of model-1

The architecture which we used in this model utilized CNN and implemented with three Convolutional layers,three max pooling layers. We used 3 x 3 kernels in the Convolutional layer & 2 x 2 kernels in the Maxpooling layers.
The images have shape of (64,64) as input. We train the convolution model with the dataset of images belongs to one class with dummy images(Elephants). Repeat the procedure for other classs of images(dogs) with dummy images (cows). Put the trained CNNs in parallel and feed them to fully connected layer having 128 neurons by using activation relu. Finally connect to softmax as the activation function. The final merged model used for compiling with sparse cross entropy with adam optimizer.
By training the merged model we get the traing accuracy 96% to 100% and validation accuracy 52% to 57%.

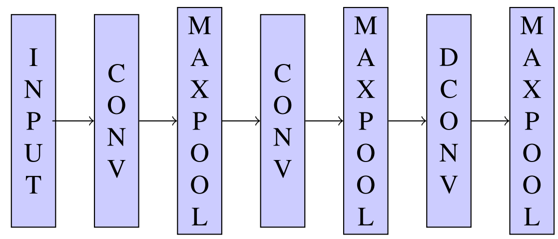

Fig. 9. Block diagram of model-2

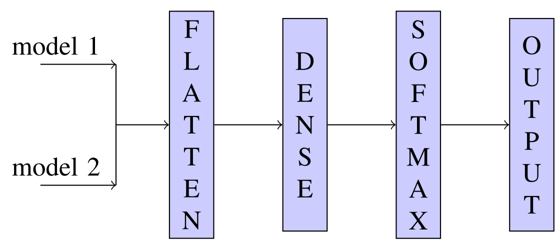

Fig. 10. Block diagram of FINAL MODEL.

C. *MULTI - CLASS IL:*

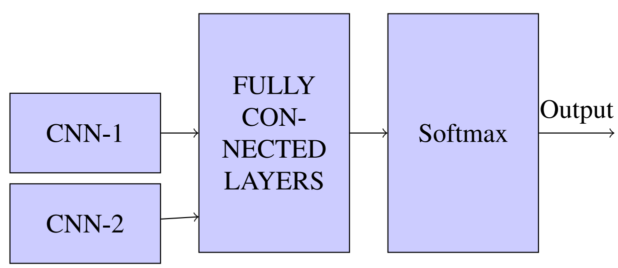

Fig. 11. Block diagram of MULTI-CLASS IL

- Train a CNN with multiple classes i.e, training phase and validation phase with good acccuracy is completed(e.g 95% accuracy).
  e.g. Multiple classes correspond to different animals : Horses, cats,..
- Extract convolutional & pooling layer outputs with freezed weights i.e. Trained CNN-1.

- Train other CNN on non living objects i.e, Extract convolutional and pooling layers with freezed weights.
- Put CNN-1, CNN-2 in parallel and feed to fully connected layers.

D. Multi and Single Class - IL:

Need for IL:

- Number of classes is unknown ahead of time.
- The trained network need not be retrained after new objects are presented to network.

E. Architecture-4:

Train the CNN for 4 classes and give it to fully connected layers.

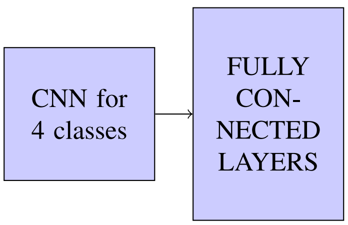

Fig. 12. Block diagram of 4-class CNN

Train the CNN for 5 classes and give it to fully connected layers.

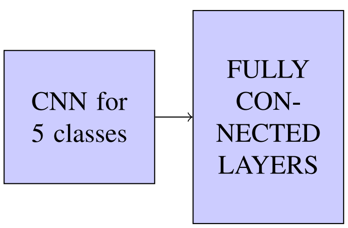

Fig. 13. Block diagram of 5-class CNN

F. Architecture-5:

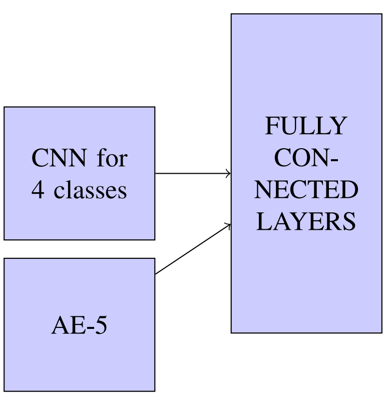

Fig. 14. Block diagram of Architecture-5

Train 4 classes with CNN (mentioned in Architecture-4) and train the 5th class with Auto-Encoder(AE). And give the output of the two corresponding models to the fully connected layers.

G. Architecture-6:

Train classes 1, 2, 3, 4 with Auto-Encoders AE-1, AE-2, AE-3, AE-4 respectively. Train the 5th class with separate Auto-Encoder. And give the output of the two corresponding models to the fully connected layers.

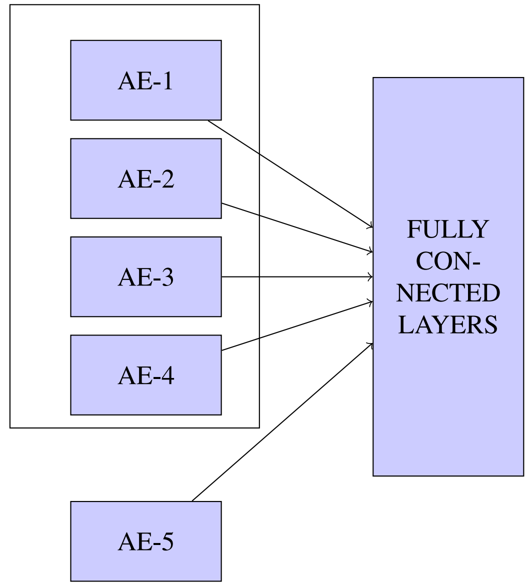

Fig. 15. Block diagram of Architecture-6

H. Architecture-7:

Train two classes (main class and dummy class) with CNN-1 followed by Auto-Encoder-1. Again train another two classes (main class and dummy class) with CNN-2 followed by Auto-Encoder-2. Put these two trained models in parallel and feed them to fully connected layers.

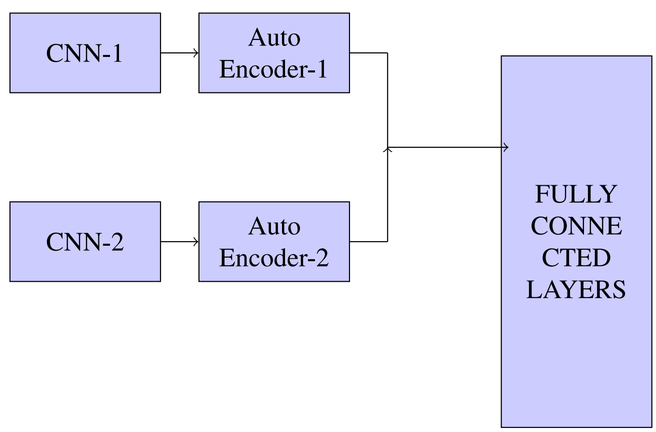

Fig. 16. Block diagram of Architecture-7

V. Experimental results:

The following table contains accuracies for different architectures with increase of number of classes.

Architectures	Training Accuracy	Validation Accuracy
Architecture-1	99%	77%
Architecture-1 for 3-classes	93%	71%
Architecture-2	88-94%	57%
Architecture-2 for 3-classes	94.33%	56.85%
Architecture-3 for 2-classes	90-97%	78.95%
Architecture-3 for 3-classes	99.34%	88.6%
Architecture-4 for 4-classes	97%	78%
Architecture-4 for 5-classes	91%	80%
Architecture-5	76.56%	69%
Architecture-6	94%	81%
Architecture-7	78.57%	78.06%

VI. CONCLUSIONS:

In this research paper using various Deep Learning architectures, Incremental Learning is demonstrated. We are actively investigating novel ANN architectures for improving the classification accuracy.

REFERENCES

[1] Deboleena Roy ,Priyadarshini Panda , Kaushik Roy , “Tree-CNN : A Hierarchial Deep Convolutional Neural Network for Incremental Learning,” Accepted in Neural Networks, 2019.

[2] Francisco M , Castro , Manuel J.Marin-Jimenez , Nicolas Guil, Cordeia Schmid , Karteek Alahari , “End-to-End Incremental Learning,” ECCV, 2018.

[3] Y.R. Asfour, G.A. Carpenter, S. Grossberg, and G.W. Lesher. (1993) ”Fusion ARTMAP: an adaptive fuzzy network for multi-channel classification”. In Proceedings of the Third International Conference on Industrial Fuzzy Control and Intelligent Systems (IFIS).

[4] B.Siva Raju, A.Sowmya and G.Rama Murthy,” Facial Emotion Detection using Deep Autoencoders,” Proceedings of International Conference on Recent Innovations in Electrical, Electronics and Communications Engineering(ICRIEECE), July 2018.

[5] B.Siva Raju, A.Sowmya and G.Rama Murthy, ” Emotion Detection using visual information with Deep Auto-Encoders, ” IEEE, 5th International Conference for Convergence in Technology,2019 (Scopus Indexed).

[6] K.I.Basha and G.Rama Murthy, ” Emotion classification: Novel Deep Learning Architectures,” Proceedings of IEEE International Conference on Advanced Computing & Communication Systems (ICACCS), 15th-16th March 2019(Available on IEEE Explore).