International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 10 (2018) pp. 8014-8017

Variational Autoencoder Coupled with Deep Generative Neural Network for the Identification of Handwritten Digits

Pallavi Saha[1*], Sreeloy Kumar Das[2], Sourav Nandy[3]

[1]Computer Science and engineering, University of Engineering and Management, Kolkata, India.

[2]Computer Science and engineering, University of Engineering and Management, Kolkata, India.

[3]Computer Science and engineering, University of Engineering and Management, Kolkata, India.

Abstract

In the present paper, a novel approach of machine learning technique has been implemented that will help the machine to learn from the experience and understand hand written digits using Deep Generative Neural Network (DGNN). The current work describes a process to generate unique images and recognize the hand written digits sent as input image to the proposed model by using unsupervised learning. In the proposed method, a variational auto encoder (VAE) has been utilized to encode and decode the input image. Inside an auto encoder, there consists of several hidden layers '*h*' which represents the input. Then clustering is undertaken and the result of clustering reflects. loss of information, so an Convolutional Neural network (CNN) lossy function is adopted to retrieve the lost information and finally DGNN is employed to predict the output and a digit generator is utilized to produce images of digits(0-9) as result of the model. Clustering of digits has been performed to show the similarities and dissimilarities among the different digits contained in the data set. Manually handwritten digits can be identified very easily but it is tough and challenging for machines to understand because handwriting differs from person to person. Although it is difficult to achieve, but it has been observed that DGNN produces acceptable results and performs identification of handwritten digits (taken input as no label data) in a better manner.

Keywords: VAE, DGNN, CNN, CLUSTERING

INTRODUCTION

Handwriting recognition is a part of pattern recognition and this becomes challenging for a machine as because a data without a label does not provide much information to the machine. It is exceptionally useful in a wide range of real life applications, including documentation analysis, mailing address interpretation, bank check processing, signature verification, document verification and many others [1]. There are many pattern recognition approaches that include statistical methods, structural and syntactic methods, and neural networks. In some cases systems identify strokes where others try to identify characters, groups of characters, or entire words. Here we have cultivated a new technique which has higher potential in better understanding digits. In our technique we have applied variational autoencoder for data compression and through deep learning using deep generative neural network. Deep learning is a part of unsupervised learning which is also a hierarchical learning that has impacted huge on the field of machine learning [2]. In this modern era Artificial neural network (ANN) has a huge contribution towards pattern recognition and machine learning. Artificial neural network are estimating system stimulated by the organic neural network that comprises in animal brain [3]. In character recognition this systems learns to identify images that contain English alphabets or digits (0-9) by analyzing example images that have been manually labelled as alphabets or digits and as a results it can be used to identify characters in other images. Since the early 1980's processing of computerized documents has been growing rapidly because of the exponentially increasing amount of daily documents. In such Documents encoding and decoding i.e. conversion of textual blocks into ASCII codes represents one of the most important tasks in document processing [4]. There are many applications of Character recognition such as Automatic Postal sorting, automatic bank cheque processing, automatic legal document verification, forensic investigation into suicidal notes, etc. Image processing as well as various machine learning techniques such as ANN, CNN, DGNN has been applied on a wide range of problems to perform identification or classification [5,6,7,8,9,10,11,12,13]. Arnold *et al.* [10] proposed a solution to the character recognition problem by employing Matlab's Neural Network Toolbox. Experimental results shows. that precision of the character recognition depends on the resolution of the character projection. This model suffered from a drawback too, it was not able to recognize the various handwriting styles. Wang *et al.* [13] put forwarded an efficient algorithm to constrain the optimization route of the visualization. Among the various inverse transformations, image blurring and deblurring has been added to the optimization process and identifiable images can be created. This algorithm reflects good results in extracting the details of the images. Brown *et al.* [14] describes features of the identification system for unconstrained handprinted symbols. The identification system uses smart thinning techniques to create centerline thinned stick figure images from raster scanned characters. The identification logic interacts with the feature extraction algorithms to extract the topological, geometrical and local measurements which will be required to recognize the character or to reject the character as unidentified.

In the current work hand written digits are recognized by the machine with the help of Variational Autoencoder coupled with Deep Generative Neural Network. The current paper is organized as follows. Section 2 discusses the various methodologies adopted. In Section 3, the proposed model has been designed to show the work flow diagram of our experiment. Section 4 reports the result of our experiment.

METHODOLOGY

Variational Autoencoder

Variational Autoencoder can be regarded as a neural network which is capable to encrypt the image into a direction in the latent space of **z** real numbers. It has been assumed that the direction to be arbitrary sample collected from a z-dimensional normal distribution. Then from the encoded vector representation, the Decoder network can then decode the vector and obtain the original image back. The reason why the latent space is z-dimensional normal distribution is that we can then draw random samples from the distribution and fetch them into Decoder network. Then we can obtain brand new images that are not even in the dataset we trained on.

The most basic VAE network links the images, each flattened from a image matrix into a single vector of length dim Then the input, of size [batch_size, dim] is fed into 3 layers of fully connected layers. Then the last fully connected layers produce a mean and standard deviation vector in the latent space. The common architecture of autoencoder is illustrated in fig. 1.

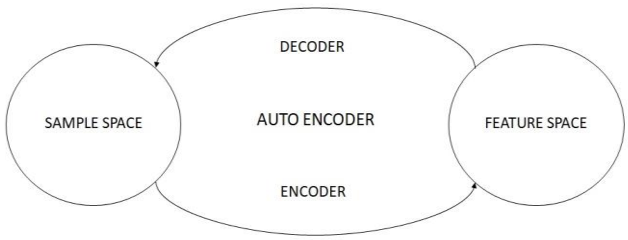

Figure 1. Represents a typical architecture of autoencoder.

Hierarchical Clustering

Hierarchical clustering is a method of cluster analysis in statistics which attempts to build clusters hierarchically. Here splitting and merging are determined in a greedy manner. The results of hierarchical clustering are represented in a dendrogram.

If a set of N items with an N x N distance (or similarity) matrix needs to be clustered, then the basic process of hierarchical clustering is mentioned below:

i. At first assign each item to its own cluster; so that if you have N items, you now have N clusters, each should contain just one item. Let the distances (with similarities) between the clusters equal to the distances (similarities) between the items they contain.

ii. Then calculate the closest (most similar) pair of clusters and merge them into a single cluster, so that we have one less cluster.

iii. Compute distances (similarities) between the new cluster and each of the old clusters.

iv. Repeat steps 2 and 3 until all items are clustered into a single cluster of size N.

CNN Loss Function

The application of the loss function is to govern the training phase of a neural network. Mean squared error and Cross Entropy loss are extensively applied for issues related to classification. In case of object detection problem, a specific type of loss function termed as focal loss is considered to facilitate the training process of the CNN based detector. A comparison is carried out by the loss layer of neural network between the ground truth, i.e., processed and reference patches, respectively and the response produced by the network.

Deep Generative Neural Network

It's the strongest way to learn in a generative distributive. manner using unsupervised learning and it has achieved tremendous success in few years. All generative models learn the true data distribution of the training set so as to generate new data points with some variations. But sometimes it becomes impossible to learn the exact distribution of data either implicitly or explicitly, so it should model a distribution which is as similar as the true data distribution. For this, we have used the advantage of neural networks to train a function which can approximate the model distribution to the true distribution.

PROPOSED SYSTEM

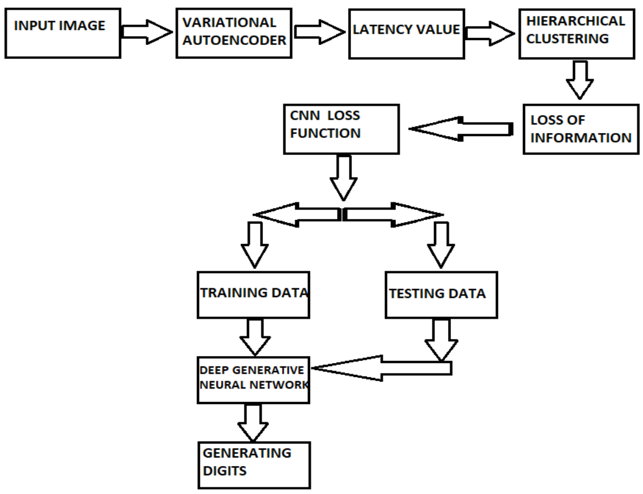

Figure 2. Describes the workflow of proposed system.

In the current work, latency values of the input images have been computed by using variational encoder. A reasonable number of instances for each type of digits are required to utilize the ANN for accurate prediction. Experiments have been conducted on a large number of handwritten digits and satisfactory results have been obtained by engaging the proposed model. Initially, a set of images (of digits) is fed as input to the variational encoder and this encoder as a result produces latency values corresponding to the input images. Then hierarchical clustering is performed on the estimated latency values. Result of the clustering process revealed that after undergoing the clustering phase a significant amount of information loss has been occurred. In order to recover the loss, the lossy function of CNN has been employed in the present work. This function has enabled to regain the lost values or information related to the input image. Once the information has been regained, the data instances are divided into two halves; former is the training phase and the latter is testing phase.

During the training phase, the deep generative neural network is skilled by using the training data. In the testing phase, an unknown input image is sent as input to the DGNN and the response produced is noted. DGNN has been utilized to predict the latency values as well as for generating the digits. The clear and concise workflow of our proposed model is illustrated in fig.2.

RESULTS AND DISCUSSION.

Simulation has been carried out on a large number of handwritten digits. Fig. 3 shows a set of input handwritten digit images. Fig. 4 depicts that digits having similar pattern are close to each other in the clustering map whereas dissimilar digits are far apart from each other. Each digits form [0- 9] are represented by different colors. The set of images for different digits generated after decoded by our model has been shown in Fig 5.

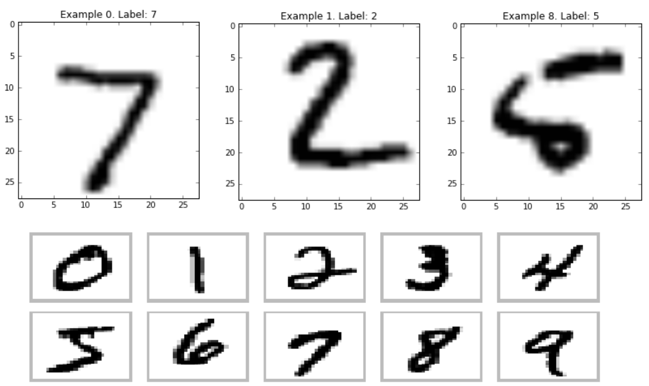

Figure 3. Depicts a set of input handwritten digit images.

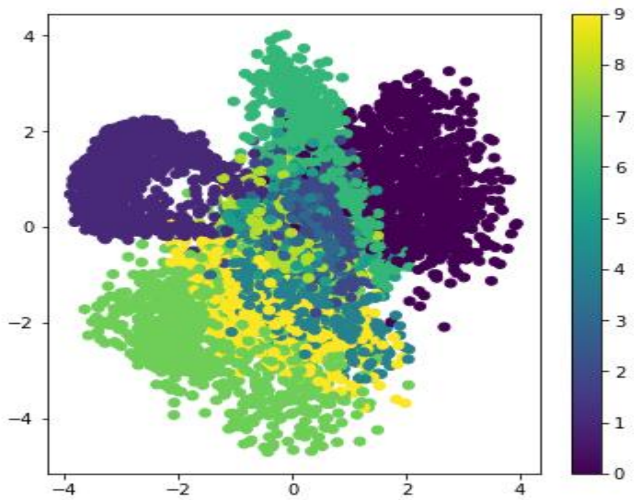

Figure 4. Results of clustering among similar and dissimilar set of images.

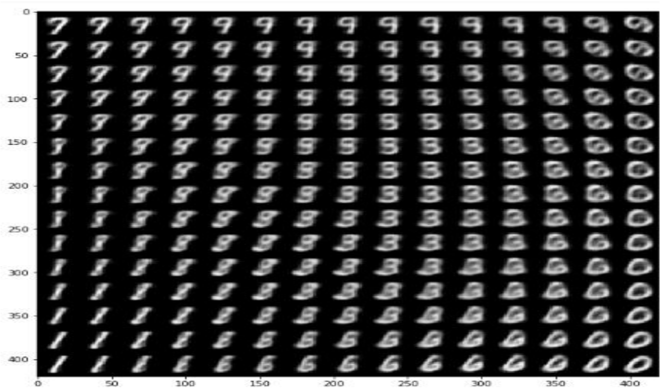

Figure 5. Digits generated after trained and decoded through our proposed model

CONCLUSION

Hand written digits are very difficult for any automated system to recognize because data without any label is very hard to understand for computers. But this problem has been eradicated by using deep learning and auto encoding techniques. Simulated results depicted that generation of the digits or prognosis of the latency values based on the input image by employing the proposed model provides better and satisfactory results.

REFERENCES

[1] Regupathi, R., Rajalakshmi, M., & Dhivya, E. (2014), “Experimental Study on Behavior of Reinforced Concrete Beams with Precast SIFCON Laminates.”, IOSRD International Journal of Engineering, 1(2), 132-139.

[2]. Li Deng and Dong Yu (2014), "Deep Learning: Methods and Applications", Foundations and Trends® in Signal Processing: Vol. 7: No. 3–4, pp 197-387.

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 10 (2018) pp. 8014-8017
© Research India Publications. http://www.ripublication.com

[3] Neural networks: the official journal of the International Neural Network Society, ISSN: 1879-2782, Vol: 61, Page: 85-117Publication Year: 2015

[4] York Manuscripts Conference, & York Centre for Medieval Studies. (1989). Latin and vernacular: studies in late-medieval texts and manuscripts proceedings of the 1987 York Manuscripts Conference. A. J. Minnis (Ed.). Brewer.

[5] M.E. Paoletti, J.M. Haut, J. Plaza, A. Plaza, “A new deep convolutional neural network for fast hyperspectral image classification”, ISPRS journal.

[6] Rosenfeld, A. and Kak, A.C., “Digital Image Processing”, Academic Press Inc., 2nd. Ed, 1993 Simon.

[7] C-L. Liu and K. Marukawa, “Normalization Ensemble for Handwritten Character Recognition”, The Ninth International Workshop on Frontiers in Handwriting Recognition (IWFHR 9), Tokyo, Japan, pp. 69-74, 2004.

[8] I. D. Jackel et.al., “A neural network approach to handprint character recognition”, IEEE Trans. PAMI, 1991.

[9] Richard Buse, Zhi-Qiang Liu and Jim Bezdek, Word Recognition using Fuzzy Logic, IEEE Trans. on Fuzzy Systems, Vol 10, No 1, Feb 2002.

[10] Rókus Arnold, Póth Miklós “Character Recognition using Neural Network”, 2010 11th International Symposium on Computational Intelligence and Informatics (CINTI).

[11] Zeiler, M., Krishnan, D., Taylor, G., Fergus, R., “Deconvolutional networks. In: IEEE Conference on Computer Vision and Pattern Recognition”. pp. 2528–2535, 2010.

[12] Jonggeol Na, Kyeongwoo Jeon, Won Bo Lee, “Toxic gas release modeling for real-time analysis using variational autoencoder with convolutional neural networks”, Chemical Engineering Science, Volume 181, 18 May 2018, Pages 68-78.

[13] Feng Wang, Haijun Liu, Jian Cheng, “Visualizing deep neural network by alternately image blurring and deblurring”, Neural Networks, Volume 97, January 2018, Pages 162-172.

[14] R.M. Brown, T.H.Fay, C.L.Walker " Hand printed symbol recognition system ", Pattern Recognition, Volume 21, Issue 2, 1988, Pages 91-118.

[15] Krizhevsky, A., Sutskever, I., Hinton, G.E., “Image net classification with deep convolutional neural networks” pp. 2392-2399.

[16] Ozan Irsoy, Ethem Alpaydin, “Unsupervised feature extraction with autoencoder trees” Neurocomputing, Volume 258, 4 October 2017, Pages 63-73.