IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-4, April 2019 E-ISSN: 2347-2693

A Hybrid System Using CNN and AE for Noisy Image Classification

Mayur Thakur [1*], Sofia K. Pillai [2]

[1,2]Dept. of Computer Science Engineering, G. H. Raisoni College of Engineering, Nagpur, Maharashtra, India

[*]*Corresponding Author: mayur.t2406@gmail.com.com, Tel.: +91-9766380096*

DOI: https://doi.org/10.26438/ijcse/v7i4.870875 | Available online at: www.ijcseonline.org

Accepted: 18/Apr/2019, Published: 30/Apr/2019

***Abstract*—** With the use of deep learning networks image processing tasks has improved due to the development of learning feature illustration from images. Generally, in the real world scenario, these images available to classify is prone to noise and other deformities. According to many types of research in the past, the deep neural networks (DNNs) are found effective for image classification problems, but they suffer from the same real-life problem of noise and other deformities in an image. Noise is common occurrences in real life situations and many studies have been carried out in the past few decades with the purpose to remove the effect of noise in the image data. In this paper, the aim was to examine the DNN-based improved noisy image classification model. We have used a hybrid of denoising autoencoder, convolutional denoising autoencoder then using a classifier which is a combination of two different architectures one is Convolutional Neural Network (CNN) and the other is extreme Gradient Boosting (XGBOOST). This technique gives progressively better outcome by incorporating CNN as a trainable element for feature extraction from the image in input and XGBoost used as an identifier at the last stage of the model for outcomes.

***Keywords*—** Blurry Images, Image Classification, Noisy Images, Supervised Classification, Unsupervised Classification, Image Denoising

I. INTRODUCTION

Development in the area of feature learning by means of deep machine learning approaches has improved the performance of image processing tasks in the last few decades, like image classification, retrieving the projected underlying information from images etc. With the capability to replicate latent features from image data using their deep layer-wise architecture gain, deep neural networks (DNN) has the potential to classify patterns. Considering smart image classification approaches, the significant research directions is the model based on neural network. Theoretically deep neural networks can estimate complex function and efficiently decipher the problems of classification. On the other hand, due to the model complication, high cost and training difficulty, this kind of structure hardly obtain very effective use. Convolutional neural network (CNN) [1-3], [4] is a ranked feature extracting variant of DNNs, which has outperformed as compared to the other image classification algorithms. The convolution layer and the pooling [5] layers keep possession of the resultant position of the features. This makes the CNN allow preserving an improved representation of the input data. There is advancement in image classification researches and still classification problem is the focus of researches because the complexity and variety of image data. CNN has been known as the most effective tool for feature extraction, however out-dated classifiers connected to CNN does not completely understand the features extracted from image. A new way out to the image classification problem is an integrated learning algorithm eXtreme Gradient Boosting (XGBoost) [6] based on GB This method achieves precise classification outcomes through iterative computation of frail classifiers. The XGBoost is applied commonly [7, 8, 9] since its high efficiency and precision.

On the other hand, execution of all such DNNs for the task of image processing particularly classification of image is inhibited because the original real life data is contaminated by different noises. DNNs require good pre-processed data for a good classification, at the time when these networks are in training. A noise result in poor classification output as the DNNs learning is affected. Several researches over image denoising have been conducted in recent years. Many studies had been performed to achieve noise reduction based on the foundation of partial differential equation centred models [10-12]. Denoising is also achieved by using wavelets as a variation of domain-transformation technique [13] and also implemented scant coding methods [14-16]. Deep architecture based mechanisms are effectively employed by the researchers for image denoising [17], [18]. CNN is used for denoising images in Jain et al. [19], which achieves superior as compared to wavelets even after utilizing a minor sample of image for training. The application of basic multi-layer perception (MLP) it is also proved analogous efficiency as compared to the earlier approaches could be attained in [20] Burger et al. Different types of model autoencoders are tested for elimination of noise from the input images and some of these approaches have extremely overtook old-style noise reduction tools. In [21] V. et al. introduced a de-noising auto-encoder (DAE) that is trained to rebuild original noise free input image from the noise damaged input noisy images and it was done via degrading the original image data with different types of noise in the course of learning period of the model. Motivated by the above proofs the main reason for the mentioned study was to create a classifier that is using supervised training. The classifier display improved performance for the noisy image classification problem and therefore explores the de-noising autoencoder (DAE) along with the convolutional de-noising autoencoder (CDAE) and incorporation of the two models i.e. CNN and XGBoost as the performance of these two, XGBoost and CNN is excellent in image classification. A hybrid of above methods is used to create a supervised classifier to improve the image classification result. Initially a DAE and a CDAE is trained in which for the output label are the noise free original images. This stage will be used to rebuild original images from noisy images at the input. For classification purpose the CNN-XGBoost is trained with unadulterated images.

The method for DAE – CDAE – CNN - XGBoost neural network configuration: The DAE, CDAE, CNN and XGBoost are cascaded in series. This design is created by using a DAE for noise elimination in fed noisy image data which is then fed to a CDAE for the purpose of improved illustration for the reassembled images and then the CNN-XGBoost model as a classifier for image classification. This approach offers superior outcome with heavy noise as the noisy input picture is being clarified using a DAE. Then this filtered image is fed to CDAE. In this stage image is further denoised and rebuilt by the CDAE. These reconstructed images after CDAE are superior fit for the classification model i.e. CNN-XGBoost classification model since CDAE rebuild the input images using additional characteristic data as it is in the original image at input and since rebuilding was achieved onto filtered reconstructed image having a reduced amount of noise. It provides more accurate output since CNN is used as trainable element for feature extraction which spontaneously provides different features and at the top level XGBoost is used as a recognizer to generate outcomes. This exceptional hybrid CNN-XGBoost model promises great consistency of classification and also features extraction. Experimental results prove of the instance that, this method outputs reasonable results and also outperform some other current methods for noisy image classification.

II. METHODOLOGY

In real world scenario imperfections like distortion other factors in image affects image classification task. Many methods based on neural network are under study for tackling noise and its effectiveness is genuinely impressive. In the following section it is explained the methods used for image rebuilding with the help of autoencoders and the classification method with CNN-XGBoost model.

2.1 Convolutional Neural Network

At the University of Toronto, Canada a group of Professors led by Yann LeCun proposed Convolutional neural network (CNN) which they applied for image acknowledgement and classification purpose of dataset consisting of manually written numeral images [22]. The network comprises a combination of eight different layers. Single input layer, dual convolution layers which achieve local features from input image and dual sub-sampling layers which is a mapping layer used for pooling operation and is also used for concluding the subordinate abstraction and 3 fully connected layers and one last output layer. The computational complexity of LeNet-5 structure of CNN is high therefore to reduce it, a simplified version is used which contains a single input layer i.e. (Input), dual Convolution Stages, dual Sub-sampling followed by the fully connected stage and at the end is the output layer (Outcome). In case of two-dimensional data the performance of CNN is great since it is a type of multi layered neural networks. In CNN, application of digital filtering method is used at each layer for the feature extraction of the witnessed data. Therefore data is transmitted over all of the several layers of the system. CNNs layer wise diagram is demonstrated in Fig.1.

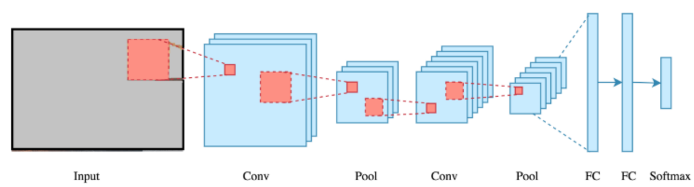

Fig.1. Convolutional neural network architecture

Representation of each numeral (from 0- 9) in the output layer can be done by certain neurons. For the specific input corresponding neuron at the output become one (1) and rest of the neuron at output stays zero (0). While training process the weights of the hidden and output layers along with systems kernel are continually improved for many epochs up to preferred precision is achieved.

2.2 De-noising Auto-encoder

Autoencoder is a type of neural network model which has 3-stage design which gives output similar as the input image. The Denoising auto-encoder (DAE) is an extension of the original auto-encoder including some stochastic additions for its ability to rebuild input original image from its distorted form. Deterministic distribution is used for manually adding noise to the input image. Initially DAE overlooks distortion in the noisy image spontaneously by way of forecasting degraded value in it. After this the DAE rebuilds the actual image by filtering the distortion. So, DAE proves to acquire the capability to create desirable feature illustration for the classifiers. Following architecture is of DAE; shown in Fig. 2.

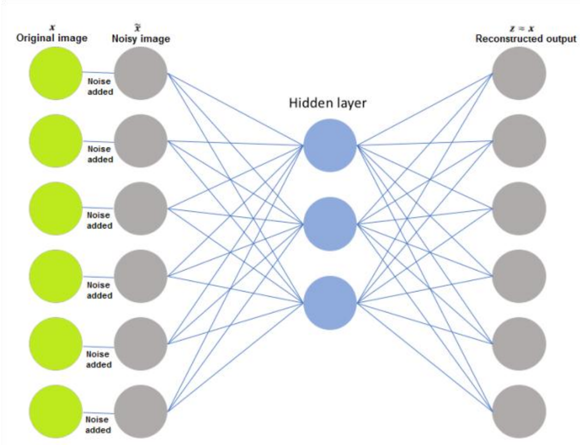

Fig.2 Denoising Auto-encoder architecture

Back-propagation (BP) is used which causes the weights between layers to update. The input to the DAE is noise added image of size 28x28 and it attempts to rebuild a disturbance free original image from it. The procedure consists of 3 steps; the input features are first encoded to hidden unit features and then rebuilds it to noise free form.

2.3 Convolutional Denoising Autoencoder

The traditional auto-encoders differ from convolutional auto-encoder (CDAE), CDAE [23] distributes the neurons weights amongst every input data's locations and by means of doing this it preserves the spatial locality. Therefore the rebuilding process is carried out through using necessary image areas in a linear arrangement which is constructed on latent code. Only one bias is used for every latent map for making all the filters dedicated on the structures of complete input image and the rebuilding is performed using back propagation method along with the de-convolution using entire shape of the image. At the times of testing period when given a noisy input in the following system it could rebuild distortion less image from it. For the output label original image data is used during the training period so that the weights of the kernel and further contributing factors can update accordingly. The architecture of the CDAE is shown in the following Fig. 3.

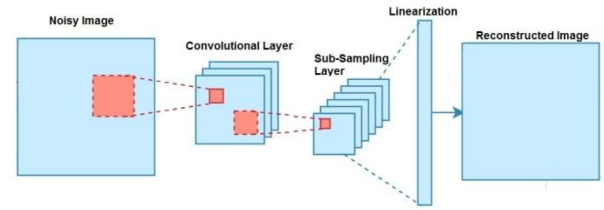

Fig. 3 Convolutional Denoising Autoencoder architecture

2.4 CNN-XGBoost Model

Initially normalization of the image at the input is done then transfers it to first layer of the CNN model i.e. the input layer. BP algorithm is used for training CNN for numerous times so that for image classification it can achieve suitable configuration. After training, the last layer of CNN i.e. the output layer which is a soft-max classifier is substituted by XGBoost and makes use of all the trainable structures of the CNN model used for training. Then above mentioned CNN-XGBoost [24] classification method gives improved classification outcomes for the testing dataset. Combining the two outstanding classifiers is useful as it can acquire features of the image at the input and delivers further accurate image classification outcomes. A CNN-XGBoost representation is given in the following Fig. 4.

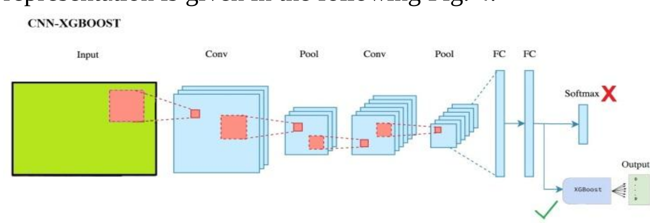

Fig. 4 CNN-XGBoost architecture

2.5 DAE-CDAE-CNN-XGBoost

The denoising and convolutional approach is combined in this proposed model DAE-CDAE-CNN-XGBoost for rebuilding original image as of noisy image. Input of DAE is fed with noisy image which is filtered and given to the next stage that is CDAE [25] [26]. The CDAE reconstruct the image removing additional noise which is then given to the final CNN-XGBoost classifier. In CNN-XGBoost the features are extracted in the CNN and XGBoost uses it to classify more accurately. The complete architecture is shown in the following Fig. 5.

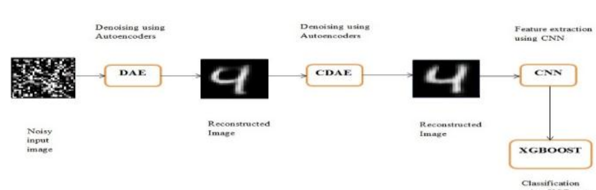

Fig. 5 Architectural block diagram of DAE-CDAE-CNN-XGBOOST

This is a three stage classifier which uses supervised learning for training. The first stage is DAE which is cascades to the second stage i.e. CDAE. Noise reduction is performed in the two stages and then lastly hybrid of CNN-XGBoost is used as an image classifier. The main purpose to implement this structure is to improve the precision for the classification of noisy image and this is done using improved reconstruction of the distorted image i.e. providing improved feature having minimum distortion. In the initial stage the weight of DAE's all layers i.e. input layer, hidden layer and the output layer are adjusted using the previously trained DAE's weight values. At stage 1 Reconstruction of the natural noise free image from its distorted form by DAE is performed and outcome generated is the reassembled image having reduced distortion as compared to the noisy image at input layer. The intermediary output which was reassembled is given at the input of stage 2 i.e. CDAE for additional reduction of distortion in image. CDAE delivers superior rebuilding as compared to DAE for noisy image problems. Since the images applied to CDAEs input are with reduced distortion as compared to the initial noisy image at stage 1 input, the model recreates superior intermediary image illustration for classification stage.

2.6 Database Description

The selected database in this paper is a very well established database for models used for image classification. MNIST is a manually written digital numeral image database that is widespread, that means it is suitable for comparative study with other methods. MNIST are handwritten 0-9 black and white image database which includes 60,000 training images and 10,000 testing pictures. Many recent studies used this dataset and also various studies considered it as its basic source of its database. The dataset has an extensive variety of diverse numeral images which accommodates 70000 sample images as per different persons rehearsing characteristic writing styles. 60,000 training images are generated by manually written numerals with the help of around 250 persons. Some images in dataset are shown in following Fig. 6.

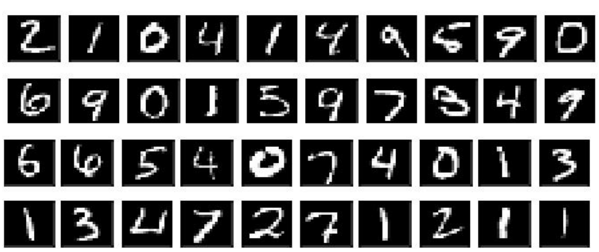

Fig. 6 Handwritten Numbers Samples (MNIST Dataset)

III. Results and Discussion

For verification about the enhancement and efficiency of the aforementioned method, comparison is carried out with the traditional ones on MNIST database. Evaluation of current method for effectiveness we compare it with other classification results of different approaches on MNIST database. The noise elimination as well as classification outcome has been examined on the dataset mentioned above. MNIST dataset i.e. 60000 samples used for training and 10000 testing images are reserved. Equal distribution of samples is done for 10 classes. This coupled insightful deep learning hybrid model DAE-CDAE-CNN-XGBoost architecture is for the heavily noisy images. The model is represented in Python Jupiter Notebook. The study has arranged on Windows 10 laptop (CPU: 8th Gen Intel Core i7 @ 1.99 GHz, GPU: 4 GB and RAM: 8 GB). Fig. 7 represents the output reconstructed image by the noise elimination at stage 1 and 2 that is DAE and CDAE respectively.

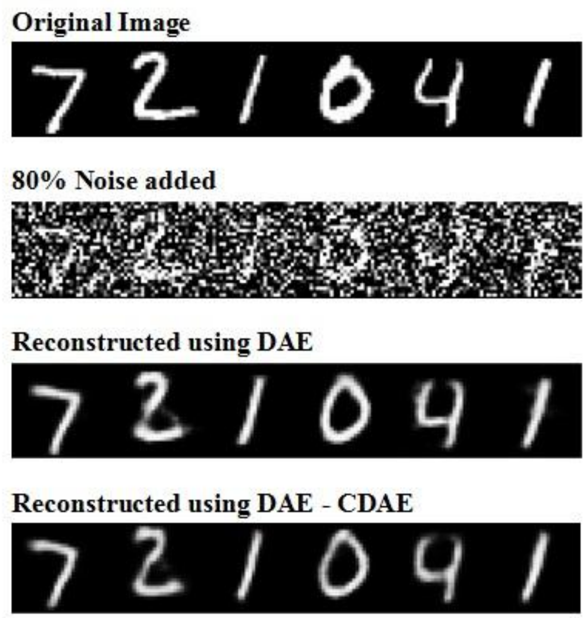

Fig.7 Examples of the output reconstructed image at stage 1 and 2

Original images of the dataset undergo pre-processing and do not contain noise. For the evaluation of the model's performance on realistic environment i.e. noisy environment, manually noise is introduced to the input image data. This is done by adding Gaussian and gamma noise with the dataset resulting corruption in the data by randomly shutting the pixels. Hybrid arrangement consisting DAE-CDAE are observed and the result clearly states that the reconstruction is improved as compared to only DAE or only CDAE. This hybrid model does not require more tuning as the two stages are already used to rebuild. At the start of training the testing dataset precision is somewhat lesser as expected but as the iteration value reached expected from these types of systems for learning, the recognition precision will improve for testing datasets rapidly. Comparison of classification's rate of accuracy of suggested model with other traditional classification models is done and model suggested in this paper has higher classification accuracy. Confusion matrix of the proposed model tested on 10,000 testing dataset examples with 80% noisy image for 200 epochs is shown in Table I. Result shown is for the proposed model in this study.

TABLE I: CONFUSION MATRIX of proposed network for 10000 test samples from MNIST dataset with 80% NOISE

DAE – CDAE – CNN-XGBOOST										
English Numeral	0	1	2	3	4	5	6	7	8	9
0	921	1	6	4	5	12	20	2	4	5
1	0	1075	6	6	6	7	4	9	19	3
2	10	14	897	40	7	3	14	24	17	6
3	0	4	24	864	2	58	4	17	16	21
4	1	8	7	4	765	16	17	18	5	141
5	12	3	7	56	4	756	18	10	14	12
6	17	6	10	2	7	27	878	2	6	3
7	1	20	19	7	12	6	0	908	1	54
8	5	10	27	88	6	71	18	17	675	57
9	6	13	4	8	74	20	0	59	7	818

IV. Conclusion and Future Scope

The real life images in actual form will consist no noise is highly impossible. Entirely studies are designed to work on pre-processed data for the classification problem this way they give good performance. In such model if the real life raw data with noise is fed to the model which is trained in supervised pre-processed data, it fails drastically. So series of auto encoder models are applied before the classification stage in such a way that the image noise reduced. This will help the classifier to improve its performance on noisy dataset. In classifiers, CNN is outstanding in performance for the image classification problem but by merging CNN model and XGBoost model it gives a classifier that delivers additional accurate results. The role of CNN is that it is used as a trainable information extractor which spontaneously acquires information from incoming images such that at the CNN models topmost level XGBoost is applied as a recognizer to generate improved outputs. The experimental results reveal that this method achieves improved result as compared with traditional CNN on the same database. There are still scopes for further developments in the architecture for better performance.

Future work should be focused on improving the CNN organization for the further extraction of the improved quality structures and speedy conjunction of the cost function. It can be done by altering the optimization methods for the improvement of the classification outcome and the effects of training.

TABLE II: Comparative description of proposed model with other methods

The work reference	Noise	Accuracy
CNN [5]	0%	98.80
	80%	11.61
CNN-XGBOOST [6]	0%	99.22
DAE-CDAE-CNN [7]	50%	96.34
	80%	85.49
Proposed DAE-CDAE-CNN-XGBOOST	80%	86.17

References

[1] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, ‘‘Gradient-based learning applied to document recognition,’’ Proc. IEEE, vol. 86, no. 11, pp. 2278--2324, Nov. 1998.

[2] A. Krizhevsky, I. Sutskever, and G. Hinton, ‘‘ImageNet classification with deep convolutional neural networks,’’ in Proc. NIPS, 2012, pp. 1097--1105

[3] J. Schmidhuber, ‘‘Deep learning in neural networks: An overview,’’ Neural Netw., vol. 61, pp. 85--117, Jan. 2015.

[4] Mayur Thakur, Prof. S. K. Pillai” A Review on Various Methods for Classification of Massive Noisy Image”, 2ndInternational Conference on Intelligent Sustainable Systems (ICISS 2019)

[5] Scherer, D., M¨ uller, A., Behnke, S.:” Evaluation of pooling operations in convolutional architectures for object recognition”. In Proc of International Conference on Artificial Neural Networks ,2010.

[6] Mackey, L., Bryan, J., Mo, M.Y.: Weighted classification cascades for optimizing discovery significance in the HIGGSML challenge. In: NIPS 2014 Workshop on High-energy Physics and Machine Learning, pp. 129–134 (2015).

[7] Schroff, F., Kalenichenko, D., Philbin, J.: Facenet: a unified embedding for face recognition and clustering. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 815–823 (2015).

[8] Song, R., Chen, S., Deng, B., Li, L.: eXtreme gradient boosting for identifying individual users across different digital devices. In: Cui, B., Zhang, N., Xu, J., Lian, X., Liu, D. (eds.) WAIM 2016, Part I. LNCS, vol. 9658, pp. 43–54. Springer, Cham (2016). doi:10.1007/978- 3-319-39937-9_4

[9] Bekkerman, R.: The present and the future of the KDD cup competition: an outsider’s perspective

[10] Perona, Pietro, and Jitendra Malik. ”Scale-space and edge detection using anisotropic diffusion.” IEEE Transactions on pattern analysis and machine intelligence , vol. 12, pp 629-639, 1990.

[11] Rudin, Leonid I., and Stanley Osher. ’’Total variation based image restoration with free local constraints.’’ , In Proc of Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference. Vol. 1. Nov, 1994

[12] Subakan, Ozlem, et al. ''Feature preserving image smoothing using a continuous mixture of tensors.'' 2007 IEEE 11th International Conference on Computer Vision. Oct. 2007.

[13] Coifman, Ronald R., and David L. Donoho. Translation-invariant denoising. Springer New York, vol. 103, pp. 125-150, 1995.

[14] Elad, Michael, and Michal Aharon. ''Image denoising via sparse and redundant representations over learned dictionaries.'' IEEE Transactions on Image processing , vol. 15, pp. 3736-3745, 2006.

[15] Olshausen, Bruno A., and David J. Field. ''Sparse coding with an overcomplete basis set: A strategy employed by V1?.'' Vision research, vol. 37 pp. 3311-3325, 1997.

[16] Mairal, Julien, et al. ''Online dictionary learning for sparse coding.'' , In Proc. of the 26th annual international conference on machine learning. ACM, pp. 689-696 June 2009.

[17] Bengio, Yoshua, et al. ''Greedy layer-wise training of deep networks.'' , In Proc of Advances in 19 th neural information processing systems, pp. 153-160, Dec 2006.

[18] Glorot, Xavier, Antoine Bordes, and Yoshua Bengio. ''Deep Sparse Rectifier Neural Networks.'' , In Proc of the Fourteenth International Conference on Artificial Intelligence and Statistics (AISTATS-11). Vol. 15. pp. 315-323, 2011.

[19] Jain, Viren, and Sebastian Seung. ''Natural image denoising with convolutional networks.'' , In Proc of Advances in Neural Information Processing Systems 21. pp. 769-776, 2008.

[20] Burger, Harold C., Christian J. Schuler, and Stefan Harmeling. ''Image denoising: Can plain neural networks compete with BM3D?.'' In Proc of Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, June 2012.

[21] Vincent, Pascal, et al. ''Extracting and composing robust features with denoising autoencoders.'', In Proc. of the 25th international conference on Machine learning. ACM, pp. 1096-1103, July 2008.

[22] Lecun, Y., Boser, B., Denker, J.S., et al.: Handwritten digit recognition with a back-propagation network. In: Advances in Neural Information Processing Systems, pp. 396–404 (1990)

[23] Masci, Jonathan, et al. ''Stacked convolutional auto-encoders for hierarchical feature extraction.'' , In Proc. International Conference on Artificial Neural Networks. Springer Berlin Heidelberg, pp. 52-59, 2011.

[24] Xudie Ren, Haonan Guo, Shenghong Li, Shilin Wang and Jianhua Li. A Novel Image Classification Method with CNN-XGBoost Model.'', C. Kraetzer et al. (Eds.): IWDW 2017, LNCS 10431, pp. 378–390, 2017.

[25] Sudipta Singha Roy, Mahtab Ahmed and M. A. H. Akhand, ''Classification of Massive Noisy Image Using Autoencoders and Convolutional Neural Network'', In 2017 8th International Conference on Information Technology (ICIT), pp. 971-979, 2017.

[26] Roy, S. S., Ahmed, M., & Akhand, M. A. H. (2018). Noisy image classification using hybrid deep learning methods. Journal of Information and Communication Technology, 17 (2), 233–269.

Authors Profile

Mr. Mayur Thakur pursed Bachelor of Engineering from Rashtrasant Tukadoji Maharaj Nagpur University in 2014. He is currently pursuing Masters in Technology. His main research work focuses on Image Processing, Machine learning and IoT

Prof. Sofia K. Pillai pursed Bachelor of Engineering and Master in Technology from Rashtrasant Tukadoji Maharaj Nagpur University. She is currently pursuing Ph.D. and currently working as Assistant Professor in Department of Computer Science Engineering, G.H. Raisoni College of Engineering, Nagpur. She is a member of IEEE & IEEE computer society. Her main research work focuses on Machine Learning, Data Mining, IoT and Artificial Intelligence. She has 11 years of teaching experience.

