IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-9, Sept. 2018 E-ISSN: 2347-2693

Effects of Varying Resolution on Performance of CNN based Image Classification: An Experimental Study

Suresh Prasad Kannojia[1], Gaurav Jaiswal[2*]

[1,2]ICT Research Lab, Department of Computer Science, University of Lucknow, Lucknow, India

[*]*Corresponding Author: gauravjais88@gmail.com, Tel.: +91-8299837473*

Available online at: www.ijcseonline.org

Accepted: 22/Sept/2018, Published: 30/Sept/2018

***Abstract*—**Convolutional neural network (CNN) based image classifiers always take input as an image, automatically learn its feature and classify into predefined output class. If input image resolution varies, then it hinders classification performance of CNN based image classifier. This paper proposes a methodology (training testing methods TOTV, TVTV) and presents the experimental study on the effects of varying resolution on CNN based image classification for standard image dataset MNIST and CIFAR10. The experimental result shows that degradation in resolution from higher to lower decreases performance score (accuracy, precision and F1 score) of CNN based Image classification.

***Keywords*—** Varying Resolution, Convolution Neural Network, Image Classification, Feature Learning, Classification

I. INTRODUCTION

Nowadays, a large amount of image data are generated and processed in real-world application [1]. Resolution of the image varies due to different input sources, different imaging devices. Variation in images resolution alters the visual information of images [2][3]. Simple visual information does not vary significantly, but complex visual information varies drastically with the reduction of image resolution. Figure 1 shows the reduction of original image resolution and their reduced visual information. The visual information plays a vital role to determine images to classify in their corresponding class.

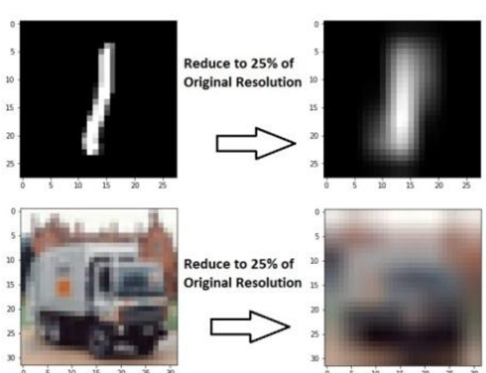

Figure 1 Reduction of resolution and visual information (images are taken from MNIST [14] and CIFAR10 [15] dataset)

With the advent of deep learning technology and the growth of computing power, convolutional neural network (CNN) has emerged one of the successful image classification models [1][4][5]. CNN based image classifier consists of the convolutional layer, pooling layer and soft-max layer. It takes input as an image, learns automatically image spatial information and preserves these spatial feature maps into their higher to lower layers [5]. Spatial-visual information constraint affects the performance of CNN based image classifier. Physical barrier to spatial visual information of an image is image quality factors (such as resolution, noise, contrast, blur, compression). In literature, most of these quality factors are experimentally visualised that how they affect the image classification performance. Dodge et al. [6] explore the effect of image quality distortions (blur, noise, contrast, JPEG and JPEG 2000 compression) on the deep neural network (VGG-16, VGG-CNN-S, GoogleNet) but not explore the effect of image resolution. While, Basu et al. [7] present modified MNIST dataset using motion blur, noise, contrast variation and successfully handle these image distortions by probabilistic quadtree DBN framework. Dejean et al. [8] show the impact of compression on CNN classification performance. They also suggest that an image can be compressed by a factor 7, 16, 40 for JPEG, JPEG2000 compression while still maintaining a correct classification. Sanchez et al. [9] analyse the impact of contrast in large-scale recognition by estimating different illumination quality. The effect of image resolution on classification is also considered. Chevalier et al. [10] propose LR-CNN model and analyse the effect of varying resolution on fine-grained classification. Ullman et al. [11] explore the effects of image resolution to classification by human and classification by DNN. Chen et al. [12] examine the effect of spatial resolution and texture window size on the performance of maximum likelihood classifier for urban land cover use. The effects of varying resolution on CNN based image classifier have not been explored in the direction of different training testing methods.

The primary goal of this paper is to visualise and analyse the experimental study of varying image resolution and its effects on the performance of CNN based image classifier. For this, an experimental methodology is proposed which have two separate training testing experiments. In the first experiment, this classifier is trained on original resolution train image dataset and evaluated on a set of varying resolution test image dataset. In another experiment, same classifier is trained on each varying resolution train image dataset and evaluated on the corresponding resolution test image dataset.

Remaining paper is organised as follow: Section II describes the experimental methodology for the study of effects of varying resolution on performance of CNN based image classifier. This methodology is implemented and evaluated on performance metrics (accuracy, precision and F1 score) [13], for standard image dataset MNIST [14] and CIFAR10 [15] in Section III. Section IV concludes the paper.

II. METHODOLOGY

The main components of proposed methodology are preparation of varying resolution images, implementation of CNN based image classifier, their training testing methods (TOTV, TVTV) and evaluation on performance metrics. The flow diagram of this methodology is shown in Figure 2.

A. Preparing Varying resolution images

For generating varying resolution sets of the original image, image rescale/resize operation is performed on an original image with a defined set of lower resolutions. For 32x32 resolution image, varying resolution image set contains 8x8, 16x16, 24x24 and 32x32 pixel resolution image. Afterwards, each image is resized into original size for the sake of input tensor of Convolutional Neural Network. Original image with varying resolution images is shown in Figure 3. In this methodology, varying resolution of standard image dataset MNIST and CIFAR10 are prepared.

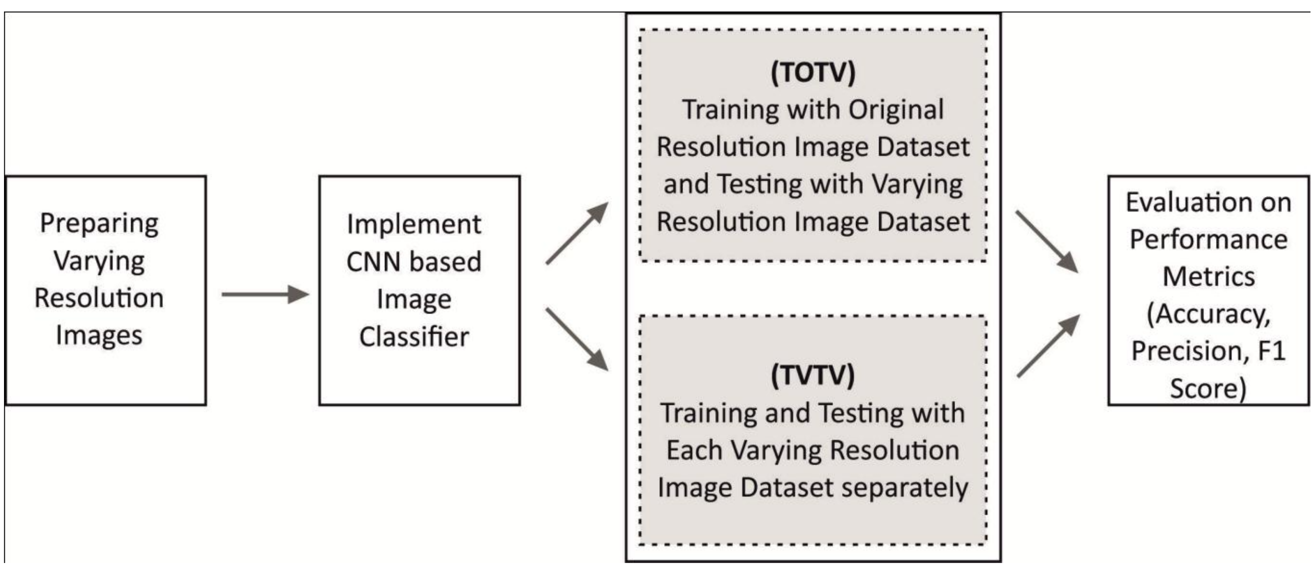

Figure 2 Flow Diagram of the Proposed Methodology

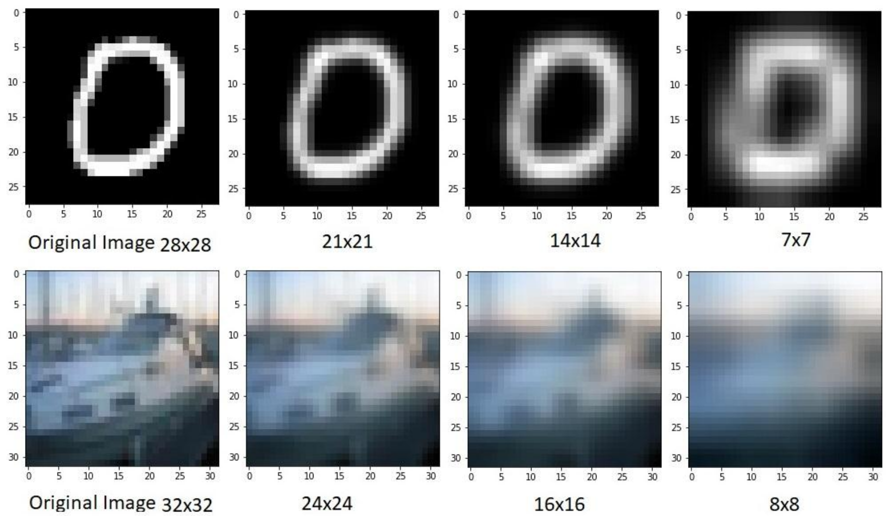

Figure 3 Original image and their varying resolution images (images are taken from MNIST [14] and CIFAR10 [15] dataset)

B. Implementation of CNN based Image Classifier

Convolutional neural network consists of mainly three types of layers: Convolutional layer, Pooling layer and Softmax layer. In convolutional layer, the input image is convolved with multiple kernels. CNN always preserve the spatial information and generate multiple feature maps. Pooling layer reduces the size of the feature map by spatial invariance average or maximum operation. Both convolutional layer and pooling layer compose feature extraction module. In the softmax layer, softmax activation function is used to classify the input feature map into class value. Traditional CNN architecture is given in Figure 4. In this methodology, CNN is implemented for image classification task. Separate CNN based image classifier is implemented for MNIST and CIFAR10 dataset respectively.

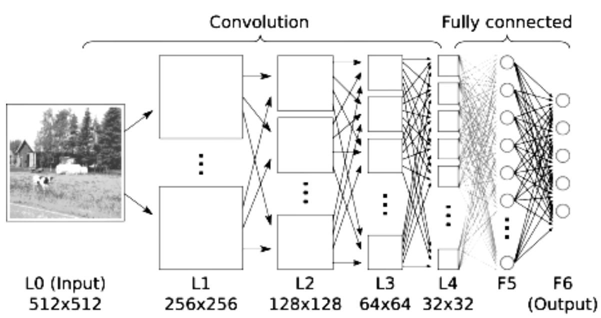

Figure 4 Traditional architecture of CNN

C. Training Testing Methods

We perform different training testing strategy on CNN based image classifier for analysing the effects of varying resolution on the performance of CNN based image classifier. From which we can evaluate the performance of learned CNN classifier on original image dataset or varying image dataset and their prediction on varying image dataset. For this, two training and testing methods are adopted. These methods are described as follows:

1) Training with original resolution image dataset and testing with varying resolution image dataset (TOTV):

This training and testing method is used to analyse how the reduction of image resolution affects the performance of classifier which trained on higher resolution images. In this method, Classifier is trained on original resolution train image dataset and evaluated on a set of varying resolution test image dataset. For 32x32 image dataset, Classifier is trained on 32x32 resolution image dataset and evaluated on separately 8x8, 16x16, 24x24 and 32x32 resolution image dataset.

2) Training and testing with each Varying resolution image dataset separately (TVTV):

This training testing method analyses the performance of CNN based image classifier for training and testing with lower resolution images. In this method, Classifiers separately trained on each varying resolution train image dataset and evaluated on the corresponding resolution test image dataset. For 32x32 image dataset, Classifier is trained separately on 8x8, 16x16, 24x24 and 32x32 resolution image dataset and evaluated on corresponding 8x8, 16x16, 24x24 and 32x32 resolution image dataset.

D. Performance Evaluation

For the performance evaluation of CNN based image classifier, standard classification performance metrics accuracy, precision, and F1 score are used in this experimental methodology. Accuracy is the fraction of correct predicted class to all predicted class. It works better for a balanced image dataset than imbalance image dataset. Precision is the ratio of number of correctly classified positive instances to the number of instances labelled by the classifier as positive. Precision metric is effective to identify actual positive from predicted positive labels whereas Recall is ratio of number of correctly positive instances to number of instances labelled are relevant. F1 score is the harmonic mean of precision and recall. Here, precision and F1 score are calculated as average per-class.

III. EXPERIMENT AND ANALYSIS

A. Experimental setup

Standard benchmark image dataset MNIST and CIFAR10 is chosen for this experimental study. MNIST image dataset contains 8 bit 28x28 resolution handwritten numerical digits (0 – 9) images. These digit images are very simple and have less visual information. CIFAR10 image dataset contains 32x32 resolution colour images. These images are complex and have high visual information. For preparing varying resolution, MNIST dataset is rescaled into 7x7, 14x14, 21x21 and 28x28 pixel resolution image dataset while CIFAR10 dataset is rescaled into 8x8, 16x16, 24x24 and 32x32 pixel resolution image dataset.

Now, CNN based image classifiers are implemented in python library sk-learn and keras with the backend of Tensorflow. Different architecture of convolutional neural network is implemented for each standard dataset MNIST and CIFAR10. Layer-wise architectural details of each CNN based image classifier is shown in Table 1. Here, both CNN based classifiers are trained and tested using TOTV, TVTV methods.

B. Result and Analysis

These experiments have been evaluated on three standard performance metrics: accuracy, precision and F1 score. The detail results of each performance score of CNN based image classifier with varying image resolution for both training testing methods (TOTV, TVTV) on MNIST and CIFAR10 dataset are shown in Table 2 and Table 3 respectively.

Table 1 Layer-wise architectural details of CNN for MNIST and CIFAR10 dataset

CNN Architecture for MNIST		
Layers	**Layers Parameter**	**Activation Function**
Conv2D	32, size=(3,3)	Relu
Conv2D	32,size=(3,3)	Relu
Maxpooling2D	Size=(2,2)	
Conv2D	64,size=(3,3)	Relu
Conv2D	64,size=(3,3)	Relu
Maxpooling2D	Size=(2,2)	
Dense	512	Relu
Dropout	0.2	
Dense	10	Softmax
CNN Architecture for CIFAR10		
Layers	**Layers Parameter**	**Activation Function**
Conv2D	32,size=(3,3)	Relu
Conv2D	32,size=(3,3)	Relu
Conv2D	32,size=(3,3)	Relu
Maxpooling2D	Size=(2,2)	
Dropout	0.25	
Conv2D	64,size=(3,3)	Relu
Conv2D	64,size=(3,3)	Relu
Conv2D	64,size=(3,3)	Relu
Maxpooling2D	Size=(2,2)	
Dropout	0.25	
Dense	512	Relu
Dropout	0.5	
Dense	10	Softmax

Table 2 Performance result of the experimental study on MNIST dataset

MNIST Dataset Resolution	**TOTV** Trained on original resolution dataset (28x28) and tested on varying resolution dataset (28x28, 21x21, 14x14, 7x7)			**TVTV** Trained and tested on each varying resolution dataset separately (28x28, 21x21, 14x14, 7x7)		
	Accuracy	**Precision**	**F1 Score**	**Accuracy**	**Precision**	**F1 Score**
28x28	0.9927	0.99269	0.99263	0.9927	0.99269	0.99263
21x21	0.9905	0.99062	0.99045	0.9924	0.99247	0.99234
14x14	0.9773	0.97828	0.97749	0.9854	0.98586	0.98544
7x7	0.6791	0.79071	0.68767	0.7770	0.84433	0.78416

Table 3 Performance result of experimental study on CIFAR10 dataset

CIFAR10 Dataset Resolution	TOTV Trained on original resolution dataset (32x32) and tested on varying resolution dataset (32x32, 24x24, 16x16, 8x8)			TVTV Trained and tested on each varying resolution dataset separately (32x32, 24x24, 16x16, 8x8)		
	Accuracy	**Precision**	**F1 Score**	**Accuracy**	**Precision**	**F1 Score**
32x32	0.8752	0.87652	0.87548	0.8752	0.87652	0.87548
24x24	0.6409	0.72365	0.65320	0.6204	0.70501	0.63220
16x16	0.3166	0.48415	0.29897	0.4233	0.62030	0.40654
8x8	0.1855	0.27090	0.13986	0.3020	0.54599	0.24262

Now, comparison graphs of each performance metrics are generated for detail analysis of the effects of varying resolution on CNN image classifier for both datasets. The performance comparison graph of classifier with varying resolution on MNIST and CIFAR10 dataset is shown in Figure 5 and Figure 6 respectively.

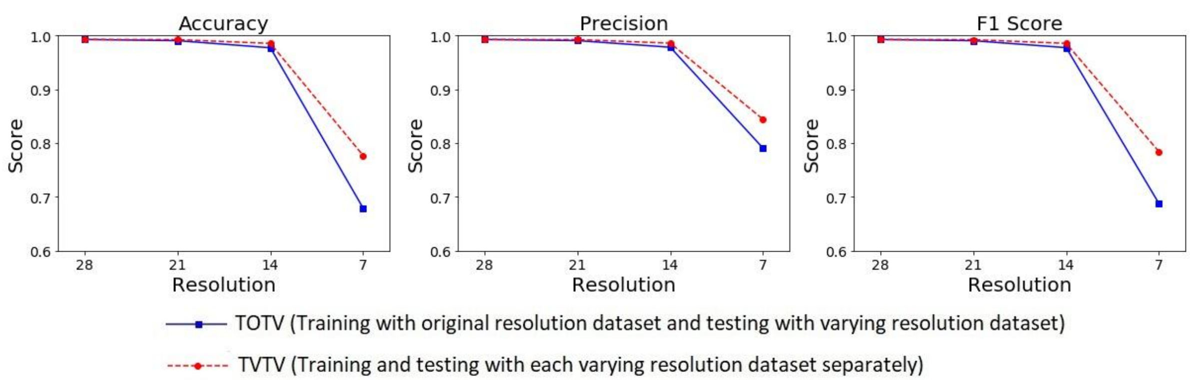

Figure 5 Comparison of performance of CNN based image classifier with varying resolution on MNIST dataset

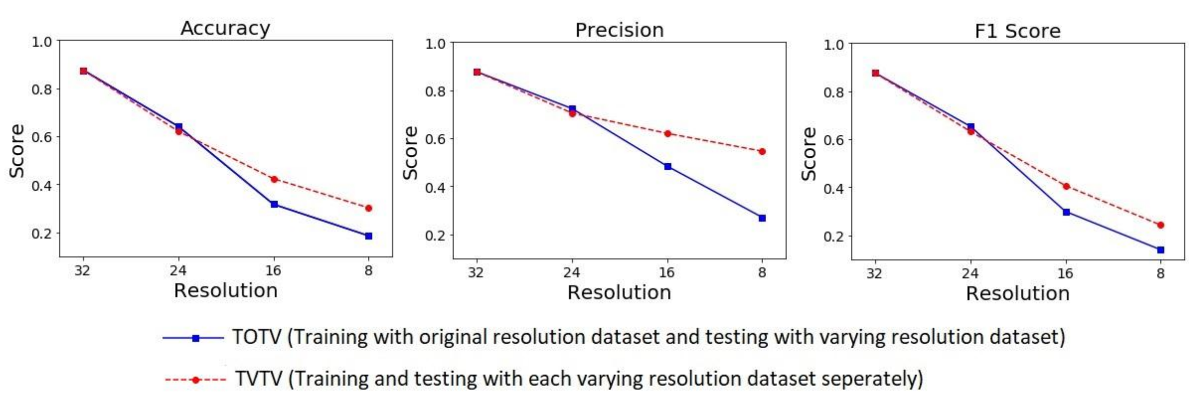

Figure 6 Comparison of performance of CNN based image classifier with varying resolution on CIFAR10 dataset

After analysing the performance comparison graph of classifier for both dataset, it is noticeable that performance score decreases when image resolution decreases. For MNIST dataset which contains images of simple visual information, the performance curve falls with little change to 14x14 pixel resolution and after this curve falls with significant change. However, for CIFAR10 dataset which contains images of complex visual information, the performance curve falls immediately with reduction of image resolution for both training testing methods. Therefore, the effects of varying resolution on the performance of classification of complex visual information images are more than simple visual information images. It is also noticeable that CNN based image classifier using TVTV training testing method is less affected than using TOTV training testing method. The precision score of both methods are higher than the accuracy and F1 score of both dataset. The higher values of precision score show that classifiers perform classification into more relevant than irrelevant images.

IV. CONCLUSION

This paper proposed a methodology and implemented on standard image datasets (MNIST, CIFAR10) for study of effects of varying resolution on performance of CNN based image classification. The experimental results and analysis conclude that performance of the classifier is mainly depended upon visual information and resolution of images. Here, degradation in image resolution from higher to lower, decreases performance score (accuracy, precision and F1 score) of CNN based image classification.

ACKNOWLEDGMENT

We are thankful to Central facility of Computational Research, University of Lucknow for providing access to KRISHNA cluster. Second author is also thankful to UGC for providing UGC-SRF fellowship to sustain his research.

REFERENCES

[1] Guo, Yanming, et al. "Deep learning for visual understanding: A review." Neurocomputing, Vol.187, pp.27-48, 2016.
[2] Sheikh, H. R., and A. C. Bovik. "Image information and visual quality.", IEEE Transactions on Image Processing,Vol.15, Issue.2, pp.430-444, 2006.
[3] Lu, Dengsheng, and Qihao Weng. "A survey of image classification methods and techniques for improving classification performance.", International Journal of Remote sensing, Vol.28, Issue.5, pp.823-870, 2007.
[4] Hoo-Chang, Shin, et al. "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning.", IEEE transactions on medical imaging, Vol.35, Issues.5, pp.1285, 2016.
[5] Deng, Li, and Dong Yu. "Deep learning: methods and applications.", Foundations and Trends® in Signal Processing Vol.7, Issue.3–4, pp.197-387, 2014.
[6] Dodge, Samuel, and Lina Karam. "Understanding how image quality affects deep neural networks.", Quality of Multimedia Experience (QoMEX), 2016 Eighth International Conference on. IEEE, 2016.
[7] Basu, Saikat, et al. "Learning sparse feature representations using probabilistic quadtrees and deep belief nets.", Neural Processing Letters, Vol.45, Issue.3, pp.855-867, 2017.
[8] Dejean-Servières, Mathieu, et al. "Study of the Impact of Standard Image Compression Techniques on Performance of Image Classification with a Convolutional Neural Network", Diss. INSA Rennes; Univ Rennes; IETR; Institut Pascal, 2017.
[9] Sanchez, Angel, et al. "Analyzing the influence of contrast in large-scale recognition of natural images.", Integrated Computer-Aided Engineering, Vol.23, Issue.3, pp.221-235, 2016.
[10] Chevalier, Marion, et al. "LR-CNN for fine-grained classification with varying resolution.", Image Processing (ICIP), 2015 IEEE International Conference on. IEEE, 2015.
[11] Ullman, Shimon, et al. "Atoms of recognition in human and computer vision.", Proceedings of the National Academy of Sciences, Vol.113, Issue.10, pp.2744-2749, 2016.
[12] Chen*, D., D. A. Stow, and P. Gong. "Examining the effect of spatial resolution and texture window size on classification accuracy: an urban environment case.", International Journal of Remote Sensing, Vol.25, Issue.11, pp.2177-2192, 2004.
[13] Sokolova, Marina, and Guy Lapalme. "A systematic analysis of performance measures for classification tasks.", Information Processing & Management, Vol.45, Issue.4,pp. 427-437,2009.
[14] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. "Gradient-based learning applied to document recognition.", *Proceedings of the IEEE*, Vol.86, Issue.11, pp.2278-2324, 1998.
[15] Krizhevsky, Alex, and Geoffrey Hinton., "Learning multiple layers of features from tiny images", Technical report, University of Toronto, Vol.1, Issue.4, 2009.

Authors Profile

Suresh Prasad Kannojia is working as Assistant Professor in Department of Computer Science, University of Lucknow, Lucknow, since 2005. He has completed Ph.D. in 2013 from University of Lucknow, Lucknow. His current area of research interest includes Pattern Recognition, Image Security, software quality, system security/reliability, data warehousing and data mining. He has also organized three national conferences and one national research scholars meet. He has published thirteen research papers in national and international journals/conferences.

Gaurav Jaiswal is Senior Research Fellow and Research Scholar in Department of Computer Science, University of Lucknow, Lucknow, India. He received his Master degree in Computer Science from BHU, Varanasi, India and Bachelor degree in Computer Science from University of Allahabad, Allahabad, India. His research interests include Computer Vision, Deep Learning, Machine Learning and Artificial intelligence.

