iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 5 **Issue:** VI **Month of publication:** June 2017

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

Volume 5 Issue VI, June 2017
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Understanding the Convolutional Neural Network & it's Research Aspects in Deep Learning

Manish Kumar Singh[1], Prof. G S Baluja[2], Dr. Dinesh Prasad Sahu[3]
Baluja Institute of Technology& Management
DTTE, The Government of NCT & Delhi
Jawaharlal Nehru University

Abstract*: *Convolutional Neural Network (CNN) is the most important Deep Neural Network (DNN) architecture to implement the Deep Learning's application of data and pattern representation in an effective and efficient manner. It uses the idea of animal's visual cortex organization to achieve connectivity pattern between its neurons. A receptive field is a restricted part of the space where a respond to stimuli are done by the individual cortical neurons. The main motivation behind the development of convolutional networks is the biological processes and CNNs are considered as the multilayer perceptrons' variations that are designed for the purpose of providing the minimal usage of pre-processing. The major applications of convolutional neural network include image recognition, natural language processing, recommender systems and video recognition. We have tried to put an honest effort in this paper to analyze the Convolutional Neural Network (CNN) and the various developments made in its area of research.

Keywords — CNN, SIANN, Receptive Fields, Neocognitron, LeNet-5, Neural Abstraction Pyramid

I. INTRODUCTION

CNN stands for Convolutional Neural Network and is the most important Deep Neural Network (DNN) architectures to implement the Deep Learning's application of data and pattern representation in an effective and efficient manner. Also referred to as ConvNet, convolutional neural Network is considered to be a kind of feed-forward ANN (Artificial Neural Network) that uses the idea of animal's visual cortex organization to achieve connectivity pattern between its neurons. A receptive field is a restricted part of the space where a respond to stimuli are done by the individual cortical neurons. The partial overlap of the myriad neurons's receptive fields results into the tiling of them to the visual field. A convolution operation is basically a mathematical approximation of the response made by the individual neurons to stimuli within the confinement of their receptive fields. The main motivation behind the development of convolutional networks is biological processes and CNNs are considered as the multilayer perceptrons' variations that are designed for the purpose of providing the minimal usage of pre-processing. The major applications of convolutional neural network include image recognition, natural language processing, recommender systems and video recognition.

SIANN acronyms for Space Invariant Artificial Neural Network and has an alternate term of "shift invariant". Both these terms of SIANN and shift invariant are used in reference to the convolutional neural network. This is due to the fact that the CNNs have their name derived from the shared weights architecture used by them and their characteristic of translation invariance.

This paper is an effort to do an analyzis of Convolutional Neural Network (CNN) concept with respect to deep learning research. The section 2 of this paper gives the overview of CNN with respect to its image recognition functionality while section 3 of this paper gives an anecdote on the developments made in the CNNs so far. The section 4 finally wraps up the paper with a suitable conclusion.

II. OVERVIEW OF CNN WITH RESPECT TO ITS IMAGE RECOGNITION FUNCTIONALITY

Convolutional Neural Network (CNN) has been modeled after following the visual perception of animal. It has been the primary area of research and application in the visual recognition field. There are several layers of receptive fields contained in CNNs. The receptive fields are collection of small neurons that are responsible for doing processing of various portions of the input image. The outputs obtained from the receptive fields are then tiled as a result of which the overlapping of input regions occur that is useful to obtain original image's representation of higher resolution. This process is iterated for every such layer. This helps CNNs in tolerating the input image's translation.

There may be the pooling layers of local or global type in the convolutional networks which are responsible for combining the neuron clusters' outputs. There also exist myriad combinations of fully connected and convolutional layers in CNNs besides the

application of point-wise nonlinearity after or at the end of each layer. On the small regions of input, the introduction of a convolutional operation lessens the amount of free parameters which is quite helpful in improving generalization. The networks have one major benefit of using shared weight in convolutional layers as a result of which the use of same filter as per pixel in the layer also lessens the memory footprint to improve the performance further.

There also exist many time delay networks which resemble the CNNs architecture, specifically those for classification or image recognition tasks. This is for this reason that neuron outputs' tiling is achievable in timed stages that are quite beneficial for doing the analyzis of images. CNNs, in comparison of other classification algorithms of images, make the usage of comparatively less pre-processing. This implies that the responsibility of the network is to learn the filters rather than hand-engineering them as they were done by the traditional algorithms. CNNs have major benefit that they are not dependent on human effort or prior knowledge at all while doing the task of feature design.

III. DEVELOPMENTS MADE IN CONVOLUTIONAL NEURAL NETWORK

The evolution of convolutional neural networks is concerned mainly with the following of visual mechanisms that are basically observed in the living organisms. The design of CNNs is the proof of how the visual techniques and complexity of living organisms are tried to be implemented in deep learning algorithm designs to help the machine to visualize the things surrounding it to distinguish the observed things to make better decision. The design of the CNNs involves various tools and techniques that are described in this section.

A. Receptive Fields

Hubel and Wiesel had presented the work in the 1950s and 1960s to show the presence of neurons in the visual cortex of the cat and monkey. They further showed that the neurons present in the visual cortex of such organisms respond individually to the visual field's small regions. This led to the evolution of receptive fields. Receptive fields are considered as those regions of visual space where the firing of a single neutron takes place after the effect of visual stimuli, provided that the eyes must not be moving. The similar and overlapping receptive field are generally found in the neighbouring cells during the firing of neuron. The size and location of receptive field varies in systematic manner in the cortex that results into the formation of a fully fledged map of visual space where the cortex lying in and around each hemisphere to represent the visual field of contra-lateral kind.

The following two fundamental types of visual cell were identified in the brain of organisms under experimentation done by Hubel and Wiesel in their 1968 paper:

1) *Simple cells:* Simple cells can have their output maximized through straight edges containing orientations of particular kind in their corresponding field.
2) *Complex cells*: Complex cells are found to contain larger receptive fields and their output is observed to be insensitive to the exact regions of the edges in the field.

B. Neocognitron

The discovery of the neocognitron took place in 1980. There is no requirement of units (located at the several positions of network) is observed in neocognitron in order to attain the same trainable weights. This idea was presented first in the book version of the corresponding original paper of backpropagation in 1986. Their development took place around 1988 for observing the temporal signals. Later in 1998, the improvisation in their design was done, their generalization was done in 2003 and their simplification was done in the same year.

C. LeNet-5

Yann LeCun et al. pioneered the LeNet. It is a 7-level convolutional network which is capable of doing the classification of digits and it applies multiple banks to do the recognition of numbers hand-written on checks which are basically 32X32 pixel digitized images. It requires bigger and larger number of convolutional layers to do the processing of images with higher resolution. However, there is a constraint associated with neocognitron network that only limited computing resources are available at present to do image processing with higher resolution.

D. Shift Invariant Neural Network

A proposal similar to LeNet-5 was made in 1988 in the form of SIANN (Shift Invariant Artificial Neural Network) for doing

character recognition of images. The modification was done in the existing architecture and training algorithm of the SIANN in 1991was made applicable in the field of medical image processing and for doing breast cancer's automatic detection in mammograms. Different design version of convolution was proposed in 1988 to find the application of doing the decomposition of electromyography convolved signals of one-dimension through the process of de-convolution. The modification in the design of such version of CNN was done in 1989.

E. Neural Abstraction Pyramid

The neural abstraction pyramid is an extended version of the feed-forward architecture of CNNs that was developed by making lateral and feedback connections in the feed-forward architecture of CNNs. This recurrent convolutional network resulted from it is capable of flexibly incorporating the contextual information for resolving the ambiguities of local kind by acting iteratively. The major benefit of this convolutional network over the previous models is that it is capable of generating image related outputs with the highest resolutions.

F. Implementations of GPU

Various publications, after following the observations of the 2005 paper where GPGPU value was established for machine learning, described myriad ways to train CNNs efficiently through the usage of GPU computing. They were later refined in 2011 in order to be implemented for impressive results on a GPU. Ciresan et al. made the significant effort to improve the performance of the multiple image databases, especially the databases of MNIST, NORB, HWDB1.0 (dataset meant for doing recognition of Chinese characters), CIFAR10 (dataset containing around 60000 RGB images of 32X32 label) and ImageNet in 2012

IV. CONCLUSIONS

Convolutional Neural Network (CNN) is the most important Deep Neural Network (DNN) architecture to implement the Deep Learning's application of data and pattern representation in an effective and efficient manner. It uses the idea of animal's visual cortex organization to achieve connectivity pattern between its neurons. A receptive field is a restricted part of the space where a respond to stimuli are done by the individual cortical neurons. The main motivation behind the development of convolutional networks is the biological processes and CNNs are considered as the multilayer perceptrons' variations that are designed for the purpose of providing the minimal usage of pre-processing. The major applications of convolutional neural network include image recognition, natural language processing, recommender systems and video recognition. CNN has become the primary area of research and application in the visual recognition field. We have analyzed the various developments made in the field of CNN including receptive field, neocognitron, LetNet-5, SIANN, Neural Abstraction Pyramid and the implementation of GPU. We can understand how promising the field of CNN and image processing. We hope the new dimensions in the field of CNN and image processing will be discovered further in future

V. ACKNOWLEDGMENT

We would like to acknowledge all the authors and scholars whose works mentioned in the references had helped us a lot to come up with this paper to analyze the convolutional neural networks and their associated developments. We would also like to say thanks to the journal – IJRASET which has given a space to our paper in it's esteem journal.

REFERENCES

[1] "Convolutional Neural Networks (LeNet) – DeepLearning 0.1 documentation". DeepLearning 0.1. LISA Lab. Retrieved 31 August 2013.
[2] Matusugu, Masakazu; Katsuhiko Mori; Yusuke Mitari; Yuji Kaneda (2003). "Subject independent facial expression recognition with robust face detection using a convolutional neural network" (PDF). Neural Networks. 16 (5): 555–559. doi:10.1016/S0893-6080(03)00115-1. Retrieved 17 November 2013.
[3] LeCun, Yann. "LeNet-5, convolutional neural networks". Retrieved 16 November 2013.
[4] Oord, Aaron; Dieleman, Sander; Schrauwen, Benjamin (2013-01-01). Burges, C. J. C.; Bottou, L.; Welling, M.; Ghahramani, Z.; Weinberger, K. Q., eds. Deep content-based music recommendation (PDF). Curran Associates, Inc. pp. 2643–2651.
[5] Collobert, Ronan; Weston, Jason (2008-01-01). "A Unified Architecture for Natural Language Processing: Deep Neural Networks with Multitask Learning". Proceedings of the 25th International Conference on Machine Learning. ICML '08. New York, NY, USA: ACM: 160–167. doi:10.1145/1390156.1390177. ISBN 978-1-60558-205-4.
[6] Zhang, Wei (1988). "Shift-invariant pattern recognition neural network and its optical architecture". Proceedings of annual conference of the Japan Society of Applied Physics.
[7] Zhang, Wei (1990). "Parallel distributed processing model with local space-invariant interconnections and its optical architecture". Applied Optics. 29 (32).
[8] Korekado, Keisuke; Morie, Takashi; Nomura, Osamu; Ando, Hiroshi; Nakano, Teppei; Matsugu, Masakazu; Iwata, Atsushi (2003). "A Convolutional Neural Network VLSI for Image Recognition Using Merged/Mixed Analog-Digital Architecture". Knowledge-Based Intelligent Information and Engineering Systems: 169–176. CiteSeerX 10.1.1.125.3812 Freely accessible.

[9] Ciresan, Dan; Ueli Meier; Jonathan Masci; Luca M. Gambardella; Jurgen Schmidhuber (2011). "Flexible, High Performance Convolutional Neural Networks for Image Classification" (PDF). Proceedings of the Twenty-Second international joint conference on Artificial Intelligence-Volume Volume Two. 2: 1237–1242. Retrieved 17 November 2013.
[10] Krizhevsky, Alex. "ImageNet Classification with Deep Convolutional Neural Networks" (PDF). Retrieved 17 November 2013.
[11] Ciresan, Dan; Meier, Ueli; Schmidhuber, Jürgen (June 2012). "Multi-column deep neural networks for image classification". 2012 IEEE Conference on Computer Vision and Pattern Recognition. New York, NY: Institute of Electrical and Electronics Engineers (IEEE): 3642–3649. arXiv:1202.2745v1 Freely accessible. doi:10.1109/CVPR.2012.6248110. ISBN 978-1-4673-1226-4. OCLC 812295155. Retrieved 2013-12-09.
[12] Le Callet, Patrick; Christian Viard-Gaudin; Dominique Barba (2006). "A Convolutional Neural Network Approach for Objective Video Quality Assessment" (PDF). IEEE Transactions on Neural Networks. 17 (5): 1316–1327. doi:10.1109/TNN.2006.879766. PMID 17001990. Retrieved 17 November 2013.
[13] Hubel, D. H.; Wiesel, T. N. (1968-03-01). "Receptive fields and functional architecture of monkey striate cortex". The Journal of Physiology. 195 (1): 215–243. doi:10.1113/jphysiol.1968.sp008455. ISSN 0022-3751. PMC 1557912 Freely accessible. PMID 4966457.
[14] LeCun, Yann; Bengio, Yoshua; Hinton, Geoffrey (2015). "Deep learning". Nature. 521 (7553): 436–444. doi:10.1038/nature14539. PMID 26017442.
[15] Fukushima, Kunihiko (1980). "Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position" (PDF). Biological Cybernetics. 36 (4): 193–202. doi:10.1007/BF00344251. PMID 7370364. Retrieved 16 November 2013.
[16] David E. Rumelhart; Geoffrey E. Hinton; Ronald J. Wiliams (1986). "Chapter 8 : Learning Internal Representations by ErrorPropagation". In Rumelhart, David E.; McClelland, James.L. Parallel Distributed Processing, Volume 1 (PDF). MIT Press. pp. 319–362. ISBN 9780262680530.
[17] Homma, Toshiteru; Les Atlas; Robert Marks II (1988). "An Artificial Neural Network for Spatio-Temporal Bipolar Patters: Application to Phoneme Classification" (PDF). Advances in Neural Information Processing Systems. 1: 31–40.
[18] LeCun, Yann; Léon Bottou; Yoshua Bengio; Patrick Haffner (1998). "Gradient-based learning applied to document recognition" (PDF). Proceedings of the IEEE. 86 (11): 2278–2324. doi:10.1109/5.726791. Retrieved October 7, 2016.
[19] S. Behnke. Hierarchical Neural Networks for Image Interpretation, volume 2766 of Lecture Notes in Computer Science. Springer, 2003.
[20] Simard, Patrice, David Steinkraus, and John C. Platt. "Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis." In ICDAR, vol. 3, pp. 958-962. 2003.
[21] Zhang, Wei (1991). "Error Back Propagation with Minimum-Entropy Weights: A Technique for Better Generalization of 2-D Shift-Invariant NNs". Proceedings of the International Joint Conference on Neural Networks.
[22] Zhang, Wei (1991). "Image processing of human corneal endothelium based on a learning network". Applied Optics. 30 (29).
[23] Zhang, Wei (1994). "Computerized detection of clustered microcalcifications in digital mammograms using a shift-invariant artificial neural network". Medical Physics. 21 (4).
[24] Daniel Graupe, Ruey Wen Liu, George S Moschytz."Applications of neural networks to medical signal processing". In Proc. 27th IEEE Decision and Control Conf., pp. 343-347, 1988.
[25] Daniel Graupe, Boris Vern, G. Gruener, Aaron Field, and Qiu Huang. "Decomposition of surface EMG signals into single fiber action potentials by means of neural network". Proc. IEEE International Symp. on Circuits and Systems , pp. 1008–1011, 1989.
[26] Qiu Huang, Daniel Graupe, Yi Fang Huang, Ruey Wen Liu."Identification of firing patterns of neuronal signals." In Proc. 28th IEEE Decision and Control Conf., pp. 266-271, 1989.
[27] Behnke, Sven (2003). Hierarchical Neural Networks for Image Interpretation (PDF). Lecture Notes in Computer Science. 2766. Springer. doi:10.1007/b11963. ISBN 978-3-540-40722-5.
[28] Dave Steinkraus; Patrice Simard; Ian Buck (2005). "Using GPUs for Machine Learning Algorithms". 12th International Conference on Document Analysis and Recognition (ICDAR 2005). pp. 1115–1119.
[29] Kumar Chellapilla; Sid Puri; Patrice Simard (2006). "High Performance Convolutional Neural Networks for Document Processing". In Lorette, Guy. Tenth International Workshop on Frontiers in Handwriting Recognition. Suvisoft.
[30] Hinton, GE; Osindero, S; Teh, YW (Jul 2006). "A fast learning algorithm for deep belief nets.". Neural computation. 18 (7): 1527–54. doi:10.1162/neco.2006.18.7.1527. PMID 16764513.
[31] Bengio, Yoshua; Lamblin, Pascal; Popovici, Dan; Larochelle, Hugo (2007). "Greedy Layer-Wise Training of Deep Networks". Advances in Neural Information Processing Systems: 153–160.
[32] Ranzato, MarcAurelio; Poultney, Christopher; Chopra, Sumit; LeCun, Yann (2007). "Efficient Learning of Sparse Representations with an Energy-Based Model" (PDF). Advances in Neural Information Processing Systems.
[33] Deng, Jia, et al. "Imagenet: A large-scale hierarchical image database."Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. IEEE, 2009.

ABOUT AUTHORS

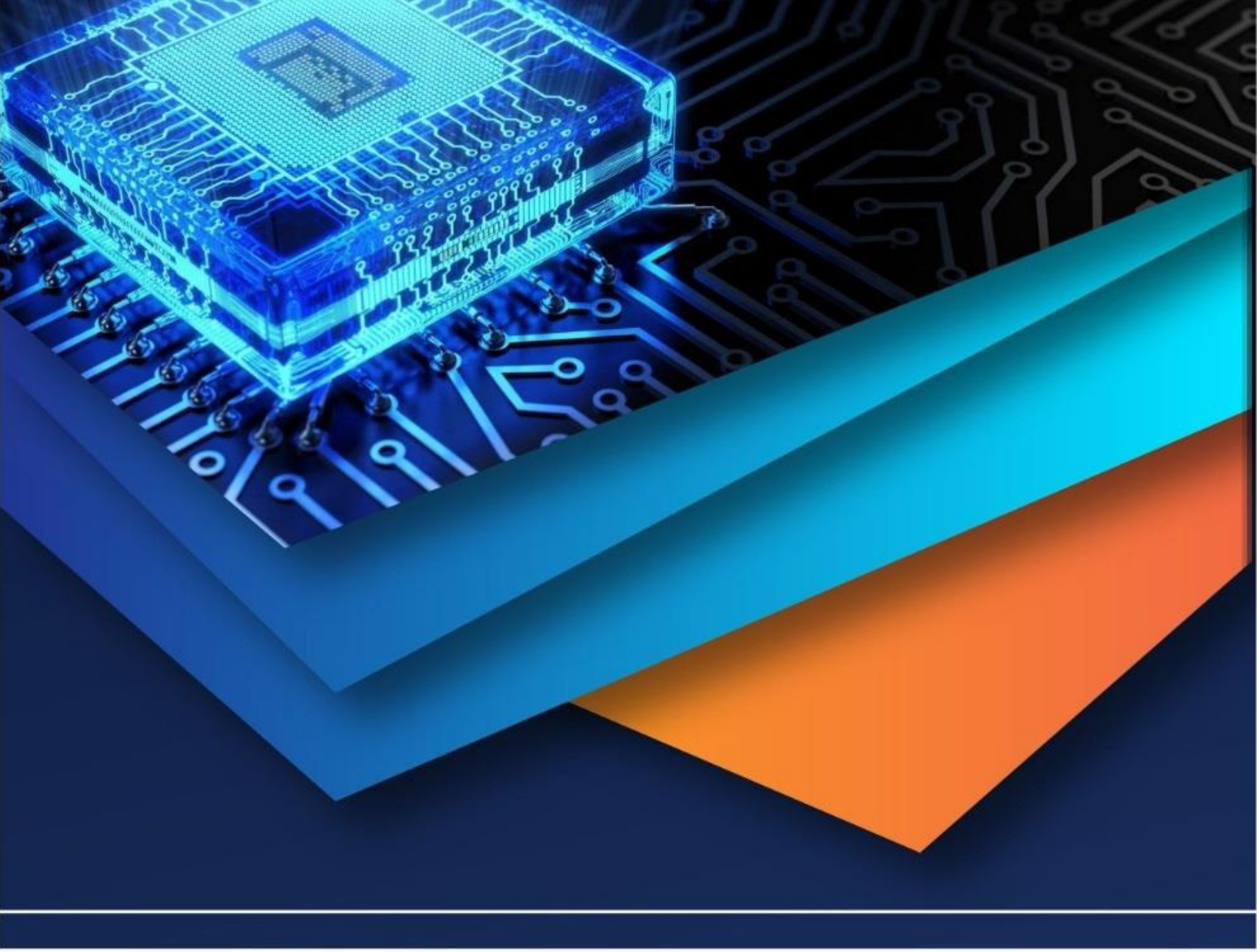
Manish Kumar has published two research papers in the international journals on web mining – chiefly “Web Mining: Penning an Era of Information Age" and “Understanding How Crucial Hidden Value Discovery In Data Warehouse Is?" and he has also

Volume 5 Issue VI, June 2017
ISSN: 2321-9653

published five articles in international journal. His areas of interest include Big Data Analytics and Machine Learning.

GS Baluja is a famous Indian author of Computer Science field who has authored numerous books so far - chiefly Data Structure Through C, Data Structures Through C++, Object Oriented Programming Using C++, Java Programming & much more. He has done B.E (Com. Sci.) from Marathwada University and M.Tech from IIT, Delhi.

Dinesh Prasad Sahu received the Master degree (Computer Science & Application) M.Tech (Computer Science & Application) from Jawaharlal Nehru University, New Delhi, India. Currently, He is doing Ph.D. (Computer Science & Engineering) under the guidance of Dr. Karan Singh, from Jawaharlal Nehru University, New Delhi, India & is working in the school of Computer & Systems Sciences, Jawaharlal Nehru University, New Delhi. His primary research interests include parallel and distributed system and Grid Computing. He has published 3 papers in proceedings of peer-reviewed Conferences.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY