JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

USE OF A FEATURE PYRAMID NETWORK FOR OBJECT DETECTION- A DEEP LEARNING TECHNOLOGY

Dr. T.A.Chavan[1], Dr.D.B.Lokhande[2],Prof.D.P.Patil[3]
[1]Principal & Professor, [2]Assitant Professor,[3]Assistant Professor
[1,2] Department of Computer Science & Engineering,[3]Electronics and Telecommunication Engg.
Shree Siddheshwar Women's College of Engineering, Solapur (MS), India

Abstract: - One of the most significant uses of deep learning technology is object recognition since it can learn features and capture them in a unique way from more conventional methods. Finding minor size differences in things is the key obstacle in object recognition. A set of multi-scale depth picture features that were taken from the backbone can be used in conjunction with top-down sampled feature pyramid networks to achieve this. Recognition systems use feature pyramids to find objects of various sizes. Pyramid representations have recently been shunned by object detectors that employ deep learning, in part because of how much memory and processing power they tend to demand. In our study, we investigate the potential of multiscale, pyramidal hierarchies to reduce the marginal costs of deep convolution networks. Top-down architecture with lateral connections is used to produce semantic feature maps with detailed information at all scales. With the feature pyramid network acting as a general feature extractor, performance gains are seen. The system is experimentally evaluated as part of the study, and the augmented pyramid network technique outperforms the FCOS model in terms of average precision (AP) by 1.2 on the MS-COCO test-dev. The study's findings showed that a feature pyramid network increased object localisation accuracy.

Keywords: FPN; Deep learning; Average Precision; Object detection; video detection;

1. INTRODUCTION

The image object detection approach examines an input image and delivers the object's category as well as its specific location. Object identification algorithms [1]–[8] have been widely employed in industry and in our daily lives as a basic task of computer vision. Without object detection technology, the roaring automatic drive, for example, cannot distinguish between nearby people, autos, or other objects. When dealing with objects of varied sizes, object detection has always been problematic. Traditional deep convolutional neural networks are extremely susceptible to changes in object scale since they aren't scale invariant. Changing an object's scale has an impact on dense object recognition. Training imbalance is a problem with the pixel regression classification algorithm. Multi-scale features are utilized in a number of ways to handle the problem of object size variation [3], [5], [9]–[14]. [3] proposed building a network of features based upon the backbone, containing features at multiple scale levels that each process objects of a particular scale range. Feature attributes at high resolution are created by upsampling low-resolution features to create high-resolution features. Classic upsampling chooses four identical locations as sampling sites around each target point, and target point attributes are produced by linearly adding the sample points' traits. In this sampling strategy, which relies exclusively on spatial correlations, points near the boundary or with other features are easily influenced by unrelated pixels. As a result, fine features are difficult to extract using only spatial coordinates while upsampling. Deep convolutional neural networks are also subjected to multiple downsampling. Upsampled features are prone to misalignment after multiple downsampling, resulting in differences and even ambiguities between restored and basic features without downsampling during fusion. Simple addition and fusion are used in the feature pyramid network. Direct addition will have an impact on the depiction of features at two levels since there are particular differences between features. Furthermore, when small item identification and accurate object localization are required, direct fusion is insufficient.

Our research presents an improved feature pyramid network to address the issues raised above. As an alternative to the core feature pyramid network for feature upsampling and multiscale feature fusion, adaptive feature upsampling (AdaUp) and feature fusion (AFF) are available. This study concludes that AdaUp is no longer based solely on spatial information, but uses semantic data as well. High-resolution low-level features1 serve as a spatial reference, while high-resolution low-level features2 help predict the offset of connected samples at each target location. Calculating the continuous coordinates of these sample sites can be done by using the offset coordinate and the coordinates of target locations. Target point attributes are created by combining the features of all sampling points via bilinear interpolation.

AnaUp allows for dynamically changing the interpolation sampling point location based on input attributes and spatial location. This makes it more adaptable than traditional interpolation upsampling. The feature fusion ratio can be dynamically changed by each pixel. High-level semantic information is necessary in other regions where judgment is a major concern, whereas low-level features are required to save detailed information in pixels in higher-detail locations. Adaptive fusion, rather than simple feature combining at two levels, may accommodate for each pixel's quirks in weight allocation, resulting in more accurate feature representation. Faster R-CNN and FCOS [8] were used as experimental benchmarks in this study to demonstrate the use of AdaFPN. Faster R-CNN was utilized in the first stage to forecast proposals with objects, and RPI Pooling [1] was employed in the second stage to collect region characteristics connected with each proposal for classification & regression.

FCOS, since it uses pixel-level categorization and regression prediction, only requires a single stage for pixel-level categorization and regression prediction. Using the MS-COCO open object identification dataset, the Faster R-CNN and FCOS models were trained and tested using FPN instead of FPN. On Faster-RCNN and FCOS, increased FPN greatly improved performance by 1.2 and 1.0 AP, respectively, based on these settings. It also outperformed its competitors in terms of localization and small-item recognition. The feature pyramid network was found to be beneficial in this study's experiments. Several experiments have been conducted in this study to confirm the validity of the feature pyramid network proposed in the proposal for object recognition and other areas of computer vision.

2. FEATURE PYRAMID NETWORK

Because of their high computational cost and lengthy hyper parameter tuning, some researchers have abandoned anchor boxes recently. By using key points and heatmap prediction, CornerNet [21], [22], [23] reduce anchor boxes, making object identification more flexible. An object box's four corners can be determined by FCOS [8] as well as each pixel can be classified based on its distance to the four corners. Because of its ease of use and efficacy, FCOS is currently frequently employed to tackle object detection difficulties in a variety of fields. Researchers used two classical methods, Faster R-CNN and FCOS, for experimental validation.

The feature pyramid network and its feature fusion block is illustrated in Figure 1. In PANet [25], new top-down, FPN-based feature maps further enhance multiscale features' representation. Using NAS-FPN, Tan et al. developed BiFPN that is more effective [26]. With feature pyramid networks, it is also possible to merge features of different sizes, as demonstrated by AugFPN [27]. [28] applied residual and dilated convolutions to the feature pyramid network to increase the feature receptive field. In order to obtain pyramidal features, CATFPN adaptively concatenates all FPN features. This is followed by downsampling and upsampling. In the proposed AFF, low-level features are suppressed by using an attention-based cross-scale fusion network, whereas in the CSFF, low-level features are also suppressed by using an attention-based cross-scale fusion network.

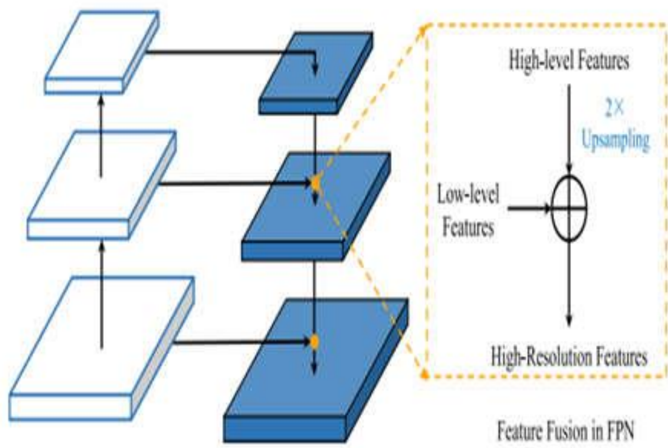


Figure 1 Feature pyramid network and its feature fusion block

3. LITERATURE REVIEW-

There are two types of image object detection algorithms: those that detect images in two stages, such as Faster R-CNN and those that detect images in a single stage, such as YOLO, SSD, and Retina Net. Pooling of ROIs [1] and alignment of ROIs [17] are commonly used in classification and fine coordinate regression. In the first stage, item suggestions are anticipated, and in the second stage, features are grouped based on propositions. Faster R-CNN [2] builds on Fast R-CNN [1] to construct object recommendations using RPN [2]. This allows us to recognize objects end-to-end. The FPN includes a pyramid network that addresses scaling concerns while also increasing the performance of faster R-CNNs.

Libra R-CNN [18] is a new class of R-CNN. By cascading several cascaded R-CNN networks together, [19] developed Cascade R-CNNs, which can improve localization accuracy over time. The pixel-level algorithms are different from two-stage algorithms because they recognize and predict things at the pixel level. Usually, anchor boxes are constructed in advance in each location and a number of them are constructed with different scales and shapes. Then, each anchor box is classified and then regressed independently. As found in [5], [20], multi-scale backbones can help identify objects in scale variation, while Retina Net[16] utilizes focal loss to help identify anchor box classification imbalances.

4. PROPOSED METHODOLOGY

In visual object detection, occlusion of objects and scale diversity are problematic. In recent years, feature pyramid networks (FPN) have been widely used to detect objects, as they create multiple-level features varying in resolution and assign to each feature varying in scale different objects. A particular scale range must be included for the capability to function at every resolution. This method of recognizing multi-level features of variable resolutions can successfully address the problem of occlusion and size variation in order to discover things.

Using a pyramid feature network, it can determine multiple levels of features, such as those found in ResNet [29]. A rudimentary feature pyramid network uses only classical interpolation and addition for upsampling features and fusing features at multiple levels. Eq. 1 can be utilized to obtain the fused feature Fo.

$$Fo = Upsample\ (Fl) + Fh\ (1) \qquad \text{------(1)}$$

A Deep Low-Resolution feature has the designation FH, while FL indicates deep low resolution features. However, because the backbone acquires image features with varying resolutions through successive downsampling and the feature pyramid network inputs multi-scale features generated from features at different layers of the backbone, reupsampling will result in feature misalignment. In this circumstance, simply integrating and combining semantic information with the structure after simple upsampling is problematic since it results in the loss of low- and high-level detail representations, as well as context information.

5. EXPERIMENTS ON OBJECT DETECTION

The COCO detection dataset [21] contains 80 categories and is tested for accuracy. To train, an 80k train picture set and a 35k val image subset (trainval 35k [2]) are utilized, and to report ablations, a 5k val image set is employed (minival). There are also final results set for the unlabeled standard test set (test-std) [21]. The training set for network backbones consists of the ImageNet1k classification set before being fine-tuned on the detection dataset. Enhanced Feature Pyramid Network is better than the baseline points in the AP of bicycle, dog, motor bike, people, bottle, boat, bus, car, cat, chair, cup, table pedestrian and mAP, respectively. Figure 2 depicts the object detection based on class condition, light condition and Indoor /Outdoor condition. Some Cityscapes detection instances, where the suggested EFPN detects more small and obstructed items. EFPN can be expanded to various situations to some extent while maintaining its accuracy enhancement.

	image	class	light	indoor/outdoor
0	2015_00001.png	1	2	1
1	2015_00002.png	1	6	2
2	2015_00003.png	1	5	2
3	2015_00004.jpg	1	3	2
4	2015_00005.jpg	1	6	2

CLASS CONDITION:

OBJECT	REPRESENTATION
"Bicycle"	1
"Bottle"	2
"Boat"	3
"Bus"	4
"Car"	5
"Cat"	6
"Chair"	7
"Cup"	8
"Dog"	9
"Motorbike"	10
"People"	11
"Table"	12

LIGHT CONDITION:

LIGHT	REPRESENTATION
"Low"	1
"Ambient"	2
"Object"	3
"Single"	4
"Weak"	5
"Strong"	6
"Screen"	7
"Window"	8
"Shadow"	9
"Twilight"	10

INDOOR/OUTDOOR CONDITION:

OBJECT	REPRESENTATION
"Indoor"	1
"Outdoor"	2

Figure 2. Object detection based on condition

EFPN can be expanded to various situations to some extent while maintaining its accuracy enhancement. By enhanced feature pyramid network, objects are detected in dark areas, on rainy days, on sunny days and on the night time based on the set of class condition, light condition and Indoor /Outdoor condition.

6. CONCLUSION

The basic feature pyramid network is improved with a novel improved feature pyramid network for feature upsampling and multi-scale feature fusion in this proposal. Object identification models Faster R-CNN and FCOS include the feature pyramid network as a recommended feature. This study describes a technique to decrease pixel-level small objects in Cityscapes that is motivated by the fact that there is a rapid increase in the number of small and obscured objects. It may be because Cityscapes and KITTI's object detection system are significantly different in terms of image quality, lighting conditions, and complexity of traffic scenes. The KITTI system has been experimentally validated with an open dataset of object detection, and it outperforms the original design significantly.

REFERENCES-

[1]. Miss. Kamble Sunayana Nivrutti, Prof. Gund V.D., et al, "Multimodal Biometrics Authentication System Using Fusion Of Fingerprint And Iris", International Journal of Trends in Scientific research and Development (IJTSRD), Sep-Oct 2018, Vol 2, Issue 6, pp 1282-1286

[2]. Kazi K. S., "Significance And Usage Of Face Recognition System", Scholarly Journal For Humanity Science And English Language, Feb-March 2017, Vol 4, Issue 20, pp 4764-4772.

[3]. T. Lin, P. Dollár, R. B. Girshick, K. He, B. Hariharan, and S. J. Belongie, ''Feature pyramid networks for object detection,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Honolulu, HI, USA, Jul. 2017, pp. 936–944.

[4]. R. B. Girshick, J. Donahue, T. Darrell, and J. Malik, ''Rich feature hierarchies for accurate object detection and semantic segmentation,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Columbus, OH, USA, Jun. 2014, pp. 580–587.

[5]. Z. Tian, C. Shen, H. Chen, and T. He, ''FCOS: A simple and strong anchorfree objectdetector,'' CoRR, vol. abs/2006.09214, pp. 1–13, Jun. 2020.

[6]. Q. Lin, J. Zhao, G. Fu, and Z. Yuan, ''Fast multi semantic pyramids via cross fusing inherent features for different-scale detection,'' IEEE Access, vol. 7, pp. 98374–98386, 2019.

[7]. Z. Guo, W. Zhang, Z. Liang, Y. Shi, and Q. Huang, ''Multi-scale object detection using feature fusion recalibration network,'' IEEE Access, vol. 8, pp. 51664–51673, 2020.

[8]. Prof. Kazi K. S., "Situation invariant Face Recognition using PCA and Feed forward Neural Networks", Proceeding of ICAEST, Feb 2016, ISBN: 978 - 81 – 930654 – 5 – 4, pp 260-263.

[9]. Prof. Nagarkar Raviraj Prakash, et al., "Pose invariant Face Recognition using Neural Networks and PCA", International Engineering Journal For Research & Development, Vol 4 special issue, pp 1-4.https://doi.org/10.17605/OSF.IO/CEVUG

[10]. Miss. A. J. Dixit, et al, "Iris Recognition by Daugman's Method", International Journal of Latest Technology in Engineering, Management & Applied Science, July 2015, Vol 4, Issue 6, pp 90-93.

[11]. Wale Anjali D., Rokade Dipali,et al, "Smart Agriculture System using IoT", International Journal of Innovative Research In Technology, 2019, Vol 5, Issue 10, pp.493-497.

[12]. Ms. Machha Babitha, C Sushma, et al, "Trends of Artificial Intelligence for online exams in education", International journal of Early Childhood special Education, 2022, Vol 14, Issue 01, pp. 2457-2463.

[13]. Pankaj R Hotkar, Vishal Kulkarni, et al, "Implementation of Low Power and area efficient carry select Adder",International Journal of Research in Engineering, Science and Management, 2019, Vol 2, Issue 4, pp. 183-184.

[14]. Karale Nikita, Jadhav Supriya, et al, "Design of Vehicle system using CAN Protocol", International Journal of Research in Applied science and Engineering Technology, 2020, Vol 8, issue V, pp. 1978-1983, http://doi.org/10.22214/ijraset.2020.5321.

[15]. Dr. J. Sirisha Devi, Mr. B. Sreedhar, et al, "A path towards child-centric Artificial Intelligence based Education", International journal of Early Childhood special Education, 2022, Vol 14, Issue 03, pp. 9915-9922.

[16]. Kutubuddin Kazi, "Lassar Methodology for Network Intrusion Detection", Scholarly Research Journal for Humanity science and English Language, 2017, Vol 4, Issue 24, pp.6853-6861.

[17]. Mr. D. Sreenivasulu, Dr. J. Sirishadevi, et al, "Implementation of Latest machine learning approaches for students Grade Prediction", International journal of Early Childhood special Education, June 2022, Vol 14, Issue 03, pp. 9887-9894.

[18]. Kazi Kutubuddin Sayyad Liyakat, Nilima S. Warhade, Rahul S. Pol, Hemlata M. Jadhav, Altaf O. Mulani, " Yarn Quality detection for Textile Industries using Image Processing", Journal Of Algebraic Statistics, July 2022, Vol 13, Issue 3, pp. 3465-3472.

[19]. Prof. Kazi K.S., Miss Argonda U A, " Review paper for design and simulation of a Patch antenna by using HFSS", International Journal of Trends in Scientific Research and Development, Jan-Feb 2018, Vol 2, issue-2, pp. 158- 160.

[20]. Ms. Yogita Shirdale, et al, "Analysis and design of Capacitive coupled wideband Microstrip antenna in C and X band: A Survey", Journal GSD-International society for green, Sustainable Engineering and Management, Nov 2014, Vol 1, issue 15, pp. 1-7.

[21]. Prof. Kazi Kutubuddin Sayyad Liyakat, "Situation Invariant face recognition using PCA and Feed Forward Neural network", Proceeding of International Conference on Advances in Engineering, Science and Technology, 2016, pp. 260- 263.

[22]. Prof. Kazi Kutubuddin Sayyad Liyakat, "An Approach on Yarn Quality Detection for Textile Industries using Image Processing", Proceeding of International Conference on Advances in Engineering, Science and Technology, 2016, pp. 325-330.

[23]. Ms. Shweta Nagare, et al., "Different Segmentation Techniques for brain tumor detection: A Survey", MM- International society for green, Sustainable Engineering and Management, Nov 2014, Vol 1, issue 14, pp.29-35.

[24]. Miss. A. J. Dixit, et al, "A Review paper on Iris Recognition", Journal GSD International society for green, Sustainable Engineering and Management, Nov 2014, Vol 1, issue 14, pp. 71-81.

[25]. Prof. Suryawanshi Rupali V., et al, "Situation Invariant face recognition using Neural Network", International Journal of Trends in Scientific research and Development (IJTSRD), May-June 2018, Vol 2, issue-4, pp. 995-998.

[26]. Ms. Shweta Nagare, et al., "An Efficient Algorithm brain tumor detection based on Segmentation and Thresholding ", Journal of Management in Manufacturing and services, Sept 2015, Vol 2, issue 17, pp.19-27.

[27]. Miss. A. J. Dixit, et al, "Iris Recognition by Daugman's Algorithm – an Efficient Approach", Journal of applied Research and Social Sciences, July 2015, Vol 2, issue 14, pp. 1-4.

[28]. Kazi K. S., Shirgan S S, " Face Recognition based on Principal Component Analysis and Feed Forward Neural Network", National Conference on Emerging trends in Engineering, Technology, Architecture, Dec 2010, pp. 250-253.

[29]. Ms. Yogita Shirdale, et al., “Coplanar capacitive coupled probe fed micro strip antenna for C and X band”, International Journal of Advanced Research in Computer and Communication Engineering, 2016, Vol 5, Issue 4, pp. 661-663.

[30]. Rahul S. Pole, Amar Deshmukh, Makarand Jadhav, et al, “ iButton Based Physical access Authorization and security system”, Journal of Algebraic Statistics, 2022, Vol 13, issue 3, pp. 3822-3829.

[31]. Dr. Kazi Kutubuddin, V A Mane, Dr K P Pardeshi, Dr. D.B Kadam, Dr. Pandyaji K K, “Development of Pose invariant Face Recognition method based on PCA and Artificial Neural Network”, Journal of Algebraic Statistics, 2022, Vol 13, issue 3, pp. 3676-3684.

[32]. Ravi Aavula, Amar Deshmukh, V A Mane, et al, “Design and Implementation of sensor and IoT based Remembrance system for closed one”, Telematique, 2022, Vol 21, Issue 1, pp. 2769- 2778.

[33]. M. Sunil Kumar, D. Ganesh et al, “Deep Convolution Neural Network based solution for detecting plan diseases”, International Journal of Pharmaceutical Negative Results, 2022, Vol 13, Issue- Special Issue 1, pp. 464-471

[34]. Dr. Kazi Kutubuddin et al , “Development of Machine Learning based Epileptic Seizureprediction using Web of Things (WoT)” , NeuroQuantology, 2022, Vol 20, Issue 8, pp. 9394- 9409

[35]. Dr. K. P. Pardeshi et al, “Implementation of Fault Detection Framework For Healthcare Monitoring System Using IoT, Sensors In Wireless Environment”, Telematique, 2022, Vol 21, Issue 1, pp. 5451 - 5460

[36]. Dr. B. D. Kadam et al, “Implementation of Carry Select Adder (CSLA) for Area, Delay and Power Minimization”, Telematique, 2022, Vol 21, Issue 1, pp. 5461 – 5474

[37]. Salunke Nikita, et al, “Announcement system in Bus”, Journal of Image Processing and Intelligent remote sensing, 2022, Vol 2, issue 6

[38]. Madhupriya Sagar Kamuni, et al, “Fruit Quality Detection using Thermometer”, Journal of Image Processing and Intelligent remote sensing, 2022, Vol 2, issue 5.

[39]. Shweta Kumtole, et al, “ Automatic wall painting robot Automatic wall painting robot”, Journal of Image Processing and Intelligent remote sensing, 2022, Vol 2, issue 6

[40]. Kadam Akansha, et al, “Email Security”, Journal of Image Processing and Intelligent remote sensing, 2022, Vol 2, issue 6

[41]. Mrunal M Kapse, et al, “Smart Grid Technology”, International Journal of Information Technology and Computer Engineering, Vol 2, Issue 6

[42]. Satpute Pratiskha Vaijnath, Mali Prajakta et al. “Smart safty Device for Women”, International Journal of Aquatic Science, 2022, Vol 13, Issue 1, pp. 556- 560

[43]. Dr. Kazi Kutubuddin Sayyad Liyakat, et al, “Voltage Sag mitigation in DVR based on Ultra capacitor”, Lambart Publications. 2022, ISBN – 978-93-91265-41-0

[44]. Dr. Kazi Kutubuddin Sayyad Liyakat, et al, “Multiple object detection and classification based on Pruning using YOLO”, Lambart Publications, 2022, ISBN – 978-93-91265-44-1

[45]. Miss. Priyanka M Tadlgi, et al, “Depression Detection”, Journal of Mental Health Issues and Behavior (JHMIB), 2022, Vol 2, Issue 6, pp. 1-7

[46]. Waghmare Maithili, et al, “Smart watch system”, International journal of information Technology and computer engineering (IJITC), 2022, Vol 2, issue 6, pp. 1- 9.

[47]. Divya Swami, et al, “Sending notification to someone missing you through smart watch”, International journal of information Technology and computer engineering (IJITC), 2022, Vol 2, issue 8, pp. 19-24

[48]. Shreya Kalmkar, Afrin, et al., “ 3D E-Commers using AR”, International journal of information Technology and computer engineering (IJITC), 2022, Vol 2, issue 6, pp. 18-27

[49]. B. Singh, M. Najibi, A. Sharma, and L. S. Davis, ''Scale normalized image pyramids with AutoFocus for object detection,'' CoRR, vol. abs/2102.05646, pp. 1–18, Feb. 2021.

[50]. J. Pang, K. Chen, J. Shi, H. Feng, W. Ouyang, and D. Lin, ''Libra R-CNN: Towards balanced learning for object detection,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Long Beach, CA, USA, Jun. 2019, pp. 821–830

[51]. G. Ghiasi, T.-Y. Lin, and Q. V. Le, ''NAS-FPN: Learning scalable feature pyramid architecture for object detection,'' in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Long Beach, CA, USA, Jun. 2019, pp. 7036–7045.

BIOGRAPHIES:

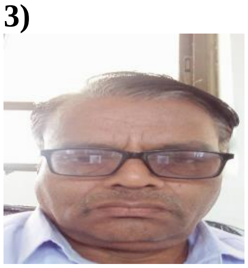
1)

Dr.Dheeraj Bhimrao Lokhande is Assistant Professor (Computer Science and Engineering) in Shree Siddheshwar Women's college of Engineering, Solapur, Maharashtra, INDIA. He is having Ph.D. in Computer Science and Engineering. He has 13 years of teaching Experience. He has qualified in GATE Examination.

2)

Dr.T.A.Chavan is Principal and Professor (Computer Science and Engg.) in Shree Siddheshwar Women's college of Engineering, Solapur, Maharashtra, INDIA. He is having Ph.D. in Computer Science and Engineering. He has 20 years of teaching Experience.

3)



Prof.Dhananjay Padmakarrao Patil is Assistant Professor (Electronics and Telecommunication Engg.) in Shree Siddheshwar Women's College of Engineering, Solapur, Maharashtra, INDIA. He has having 17 years of teaching experience and 6 years of industry experience.