International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-12, October 2019

Intelligent Coding Unit Partitioning using Predictive Data Mining

Chhaya Shishir Pawar, Sudhir Deoraoji Sawarkar

Abstract: Increasing applications of videos in everyday life demands compressing the videos further. International bodies for Video Coding standards are working toward making it more efficient in terms of reducing bitrate so as to efficiently compress the high-resolution videos. With increasing resolution, the size of the Coding Units increases. Latest Video Coding techniques like High Efficiency Video Coding (HEVC) and Versatile Video coding (VVC) proposed Larger coding Units with flexible Quadtree decompositions. In Inter-picture prediction all the sub blocks have to find best partitioning structure during motion estimation. Due to larger coding units finding the best partitioning introduces computational complexity. In the proposed work we present a computational complexity control scheme using predictive data mining. The method helps to predict whether to split or no split the coding unit. The decision tree model trained offline in the proposed work achieves 77.73% saving in encoding time with minimal change of 0.15 in average PSNR and 0.00074 in average SSIM values.

Index Terms: CU partitioning, HEVC Inter-prediction, predictive data mining, Video Coding.

I. INTRODUCTION

In today's era videos finds its application in every facet of our life. Right from social media, video conferencing, Gaming, Medicine, Education, Video calling, Television, Streaming media applications, telephony etc. Joint collaborative team on video coding released a video coding standard called High Efficiency Video Coding in 2013 which reduces the bitrate to half as compared to H.264, the currently dominating standard of the market today.

HEVC introduces the Concept of Coding tree Unit of size 64x64 which further divide into flexible CU sizes from 64x64 to 8x8. The increased block sizes help to cope up with growing resolution of the videos. HEVC also suggests the use of flexible partitioning structure. Every coding unit undergo quadtree decomposition to find the best matching block for motion estimation process. The process puts lot of burden on the encoder and consumes majority of computational time

Video is generally divided into I frames and P frames. I frame are spatially predicted i.e. without reference to any other frame but only with reference to its own blocks of pixels. whereas P frames are predicted using temporal information i.e. they are predicted with reference to previously coded frames. Only the difference from the previously coded frame is sent in the form of motion vectors. Due to larger coding Unit sizes in latest video encoding techniques and their flexible quadtree partitioning structure the complexity of the encoder increases dramatically. Hence it puts lot of burden on the encoder but assures the high efficiency in bitrate reduction.

In the proposed work we demonstrate how we can speed up the CU split decision process and control the complexity of the encoder with the help of predictive data mining technique. These techniques are useful in gaining knowledge from the previously encoded data from the training set and predicts the splitting decision at run time.

II. CODING UNIT PARTITIONING STRUCTURE

Increased resolution has compelled the use of larger size block sizes. H.264 has block maximum block size of 16x16. Whereas HEVC allows the largest coding unit of size 64x64 till 8x8. It also makes use of flexible partitioning structure. Both these features make the final Coding unit partitioning very much content adaptive in nature resulting fine details being captured and maintaining the quality of picture.

HEVC provides 50% bitrate saving as compared to its predecessor. More than half of the bit rate savings are result of its flexible partitioning structure. Larger block sizes are necessary to handle higher resolution content but at the same time smaller block sizes are required to take care of the local characteristic in the picture.

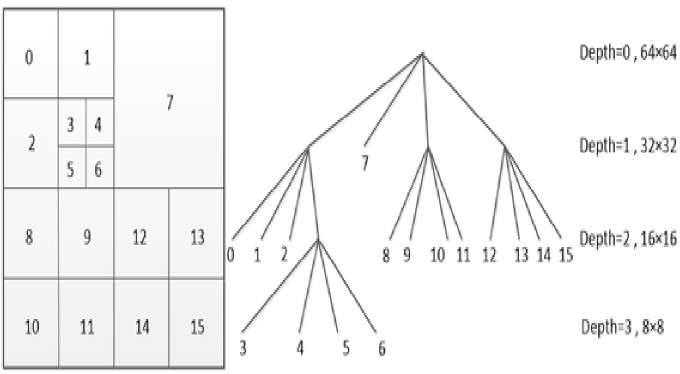

Fig1: Block Structure of HEVC

One of the main contributions of H.265 is its block partitioning structure. Each picture is divided into square shaped blocks also known as coding tree blocks.

Revised Manuscript Received on October 30, 2019.
***** Correspondence Author
Chhaya Shishir Pawar*, Department of Computer Engineering, Datta Meghe College of Engineering, Navi Mumbai, India.
Dr.Sudhir Deoraoji Sawarkar, Department Of Computer Engineering, Datta Meghe College of Engineering, Navi Mumbai, India.
This work is supported by Minor research grant from University of Mumbai

Every CTB is further divided into recursive quadtree blocks of smaller These larger sized CTUs are good for high resolution but brings more memory requirements, Encoder and decoder delay and most importantly huge computational complexity burden.

Multiple coding units together form Coding Tree Block. These CUs take variable sizes in the form of PU i.e. Prediction Units, TU i.e. Transform Units for the purpose of Intra picture prediction, motion compensated prediction and transform coding. Different applications can select the suitable sizes as per their requirement while maintaining the quality and complexity trade off.

Due to the Quadtree decomposition of the CTUs every smaller block of all the sizes right from 64x64 till 8x8 has to undergo the searching for best prediction during motion compensation process during inter-prediction and intra-prediction. Hence though this feature contributes to reduction of bitrate but brings lot of complexity for the encoder. The increased complexity of the encoder is motivation behind this work and also many researchers are working toward controlling this encoder complexity and at the same time not compromising on the quality.

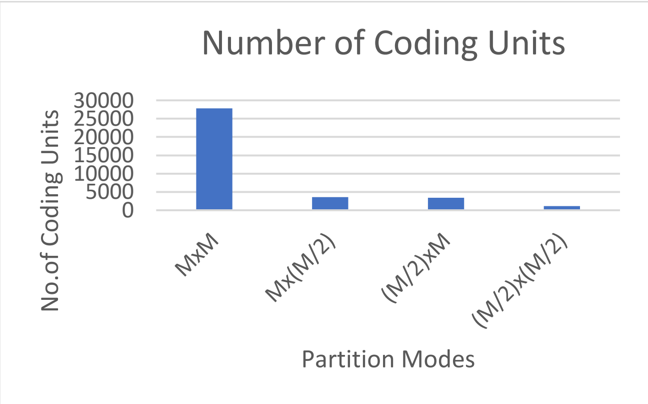

Fig 2: Coding Unit mode distribution

As we can observe from Fig 2 Maximum number of Coding units does not require quadtree partitioning as they form the part of homogeneous background. In natural scenes homogeneous background is more often than other partition modes.But still the encoder wastes a lot of computational efforts in calculating the prediction of all the CUs for all the modes of splitting. Hence it is justifiable to look for a solution that saves these efforts or can directly predict the partition mode of the coding unit.

III. LITERATURE SURVEY

Kalyan Goswami et al (2015) [1] proposed a fast method for high efficiency video coding which makes use of markov Chain Monte Karlo model for initial training of the Bayesian classifier and then Bayesian classifier during run time updates its own prior probability values and conditional values. The Bayesian classifier is used to decide the Coding unit termination.

Mai Xu et al (2018) [2] presents a Convolution neural network and long and short memory-based approach to solve the computational complexity arising out of recursive quadtree splitting structure. Initially a large training database of coding unit data is collected for deep learning of the problem and then learned network assists in early CU termination.

Yang Xian et al (2018) [3] designed a multi-level machine learning based solution for VP9 coding unit partitioning structure. They have used a weighted support vector machine to decide whether to terminate the quadtree splitting of blocks or to continue till it gets the best prediction.

Kalyan Goswami et al (2014) [4] suggests the use of ratio function of RD costs of present coding units to its next depth level units. And the ration helps deciding the splitting decision of the CU.

Buddhiprabha Erabadd et al (2018) [5] proposed a support vector machine-based solution to complexity reduction. They studied the online as well as offline model training of SVM and reported that off line training performs better in terms of time saving.

Bochuan Du et al (2015) [6] worked on random forest-based method for early CU termination method. They have studied the partitioning depth information from neighboring I.e. left and top coding units to decide the depth of the current coding unit.

Mateus Grellert et al (2018) [7] used a support vector machine-based approach which predicts whether to continue with exhaustive search for the best prediction or to early terminate the Partitioning procedure. The authors performed the experiments for various bitrates.

Linwei Zhu et al (2017) [8] demonstrated the use of Support vector machine-based solution for CU depth level selection as well as PU depth level selection. It makes use of two classifiers and multiple reviewer system to perform voting among the classifiers.

Xin Lu et al (2017) [9] proposes a CU depth decision scheme based on DCT energy of largest coding units. Optimal CU partitioning structure is obtained by setting adaptive threshold on AC energy of LCUs Yih-Chuan Lin et al (2014) [10] used the sobel operators for edge detection within the LCU and if the intensity of the edges goes beyond the threshold value then accordingly the coding unit depth decision is taken.

Svetislav Momcilovic et al (2015) [11] presented a run time dynamic neural network which does not require any training during run time. It dynamically adapts to scene content of the video.

Chan-seobPark et al (2014) [12] uses the contextual information from forward and backward frame also the coding unit from the next depth level to decide for merge skip decision process in quadtree partitioning structure.

Sangsoo Ahn et al (2014) [13] proposes a system making use of spatial information n the form of SAO parameters and temporal parameters to compute the texture characteristics of the coding units to decide the splitting of blocks.

Yun Zhang et al (2015) [14] suggests the use of multiple classifier and perform the voting of prediction output. It reduces the time required for coding unit partitioning. It makes use of support machine.

IV. THEORETICAL BACKGROUND OF DECISION TREE CLASSIFIER

Predictive data mining is capable of giving solutions to almost every problem where large amount of data is available for analysis. The huge amount from data available from the picture in terms of prediction units and dominant use of particular mode motivated us to make use of data mining technique. Classification techniques of data mining are helpful in intelligent decision making. Decision trees are the optimum choice as We would not like to overburden the encoder further. Decision trees can be implemented in the form of simple if-else statements which are computationally inexpensive. These are used to predict whether to split or not to split the coding unit.

Decision tree induction is constructing decision trees from the class labeled training dataset. Decision tree consist of tree like structure which originates at root node and each internal node represent the attributes used for classification. The leaf nodes at the end of the branch resent the final outcome in the form of predicted class label.

Attribute selection measure is the criteria for selection of splitting attribute at the nodes. It gives ranking to the attributes as per the highest value of attribute selection measure. One of the measures is Gini Index. Gini index is used for measuring the impurity of D which is a training dataset.

$$Gini(D) = 1 - \sum_{i=1}^{m} p_i^2,$$

The probability of the tuple to belong to classs Ci is given by piand it is calculated as |Ci|/|D|.The sum is calculated over number of classes . The reduction in impurity due to split on attribute A is

$$\Delta Gini(A) = Gini(D) - Gini_A(D).$$

The attribute with maximum reduction in impurity or minimum Gini index values is taken as splitting attribute.

V. PROPOSED METHODOLOGY

Every frame is divided into coding units which int turn take variable sizes for prediction units which is used in motion compensation process. While doing this in order to find if the current coding unit is to be divided further. If the coding unit is part of homogeneous contents then its not required to be decomposed else it requires dividing symmetrically or asymmetrically into prediction units.

The proposed work presents a predictive data mining-based solution which assists the encoder in deciding whether the coding unit is to be split or to continue with the motion compensation with the block as a whole.

The proposed methodology presents a low complexity model for speeding up Intra-picture prediction in High efficiency Video coding. The model consists of decision tree which is trained offline. The trained model predicts the most suitable mode during run time. The method gives substantial encoding time saving during run time with minimal loss of Peak Signal to Noise ratio and Structural Similarity Index.

The model proposed here is a low complexity model because introduction of data mining module for prediction of CU split decision will introduce additional computation. That must outperform the time required for brute force method of splitting decision. The following factors make the CU depth decision intelligent.

- Use of simpler predictive data mining model like decision tree which is easy to build, understand and the model is fast at prediction during run time because these are essentially if-else statements requiring nominal computation time.

- Features used SSD (Sum of squared difference), SAD (Sum of Absolute Difference), MSE (Mean square Error), Correlation, STDV (Standard Deviation), MotionX (Motion Vector in X direction), MotionY (Motion Vector in Y direction), RD cost. These parameters are calculated between Current MxM Coding unit block and the corresponding MxM Coding Unit in the previously reconstructed frame.

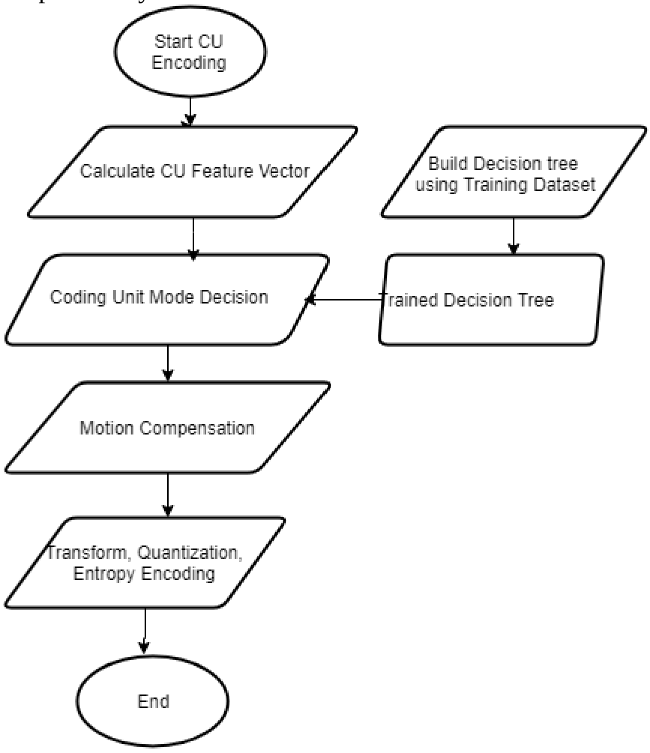

Fig 3: Flowchart of the Proposed Method

As shown in fig. 3 All these features require very minimal computing requirements. These are calculated for MxM size coding unit. Based on the values of the feature vector the trained model predicts whether to split the coding unit further or not. It predicts the partitioning mode of the CU quickly. Thereby avoiding the recursive calls to mode decision process for coding unit from 64x64 to 8x8.

VI. RESULTS AND DISCUSSION

The proposed method models the CU splitting decision problem as a multiclass classification. The Classifier will predict the mode of the Coding Unit decomposition. The Training of the classifier is done in Matlab Environment. The *Akiyo, Bus, Husky, Bowing, City* are the test sequences used for the training the classifier. The training sequences are selected so that they cover different types of motion, level of detail in the frames, and mixture of homogeneous and non-homogeneous contents.

We have run the experiment for Quantization parameter value 27 and block size of 16x16.

The metrics used for evaluation includes Encoding Time, Peak Signal to Noise ratio and Structural Similarity Index. All the parameters were verified against the standard HEVC CU splitting brute force method.

% Encoding Time Saving=100-($ET_{proposed/}$ $ET_{standard}$*100)

Where $ET_{proposed\ is}$ the Encoding time required by proposed method and $ET_{standard}$ encoding time required by the HEVC standard method. Difference in PSNR is given by

$$\Delta\ PSNR=PSNR_{standard}-PSNR_{proposed}$$

where $PSNR_{standard}$ and $PSNR_{proposed\ are}$ the PSNR values for standard method and proposed methods respectively.

$$\Delta\ SSIM=SSIM_{standard}-SSIM_{proposed}$$

where $SSIM_{standard}$ and $SSIM_{proposed\ are}$ the SSIM values for HEVC standard method and proposed methods respectively

Table 1: Performance of the Proposed Method

Test Sequence	% Saving in Encoding Time	Δ PSNR	Δ SSIM
Mobile	78.9488	0.071	0.00057
Silent	81.9060	0.090	0.00050
Ice	79.4080	0.170	0.00210
Grandma	82.0147	0.090	0.00030
Paris	75.8886	0.060	0.00030
Highway	79.5754	0.210	0.00120
Harbour	79.6091	0.060	0.00030
Claire	71.7152	0.500	0.00080
coastguard	70.5152	0.100	0.00060
Average	**77.7312**	**0.150**	**0.00074**

Table1 shows the encoding time saving done by the proposed method. It almost saves 77% encoding time as compared to original exhaustive method of finding the best partitioning structure. Substantial encoding time saving per test sequence can be observed in Fig.4.

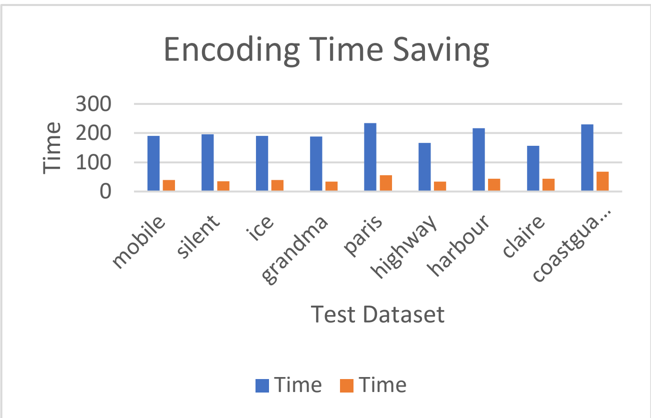

Fig 5: Encoding Time Saving by Proposed Method

Fig.6 and 7 shows the minimal loss in PSNR and SSIM values which outperform the coding gain that we get by the proposed method. The decision tree classifier with low complexity features works so efficiently in terms of time saving that the loss in PSNR and SSIM is very much negligible.

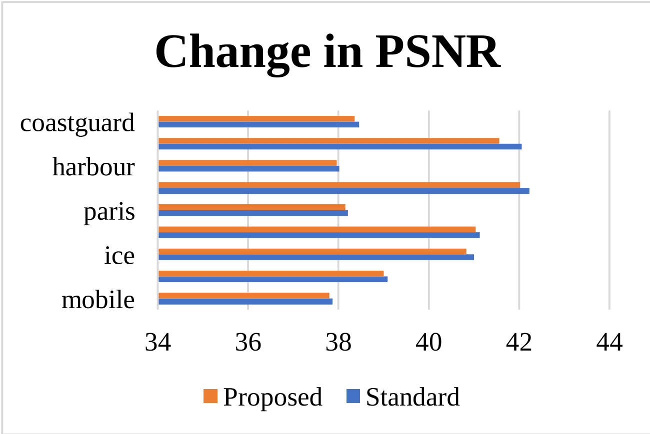

Fig.6: Minimal Loss in PSNR by Proposed Method

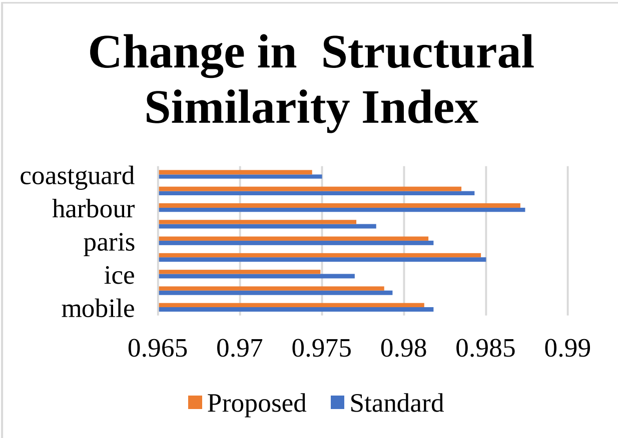

Fig.7: Minimal Loss in SSIM by Proposed Method

VII. CONCLUSION

We presented a predictive data mining-based technique which learns using the huge amount of data generated by the training dataset and predicts the Coding Unit partitioning mode very well in terms of time saving with almost no loss in quality. It overcomes the computational burden caused by the larger coding units and flexible partitioning structure.

The proposed method is intelligent as it learns from the training data and quickly predicts in run time for the current Coding Unit avoiding the wastage of computational efforts and still not compromising on the quality much.

The method speeds up the inter prediction phase of the high efficiency video coding which contributes to majority part of video coding. The proposed method will be very useful not only for HEVC but also for VVC i.e. Versatile video coding in future as the coding block sizes are still higher and hence there is lot of scope for encoding time saving as well.

REFERENCES

1. Kalyan Goswami and Byung-Gyu Kim," A Design of Fast High Efficiency Video Coding (HEVC) Scheme Based on Markov Chain Monte Carlo Model and Bayesian Classifier" Journal Of L Atex Class Files, Vol. 14, No. 8, August 2015.
2. Mai Xu, Tianyi Li, Zulin Wang, Xin Deng, Ren Yang and Zhenyu Guan "Reducing Complexity of HEVC: A Deep Learning Approach" IEEE Transactions on Image Processing (Vol.27, Issue:10,Oct 2018) H. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1985, ch. 4.
3. Yang Xian. In, Yunqing Wang, Yingli Tian," Multi-LevelMachineLearning-basedEarlyTerminationinVP9 Partition Search" IS&T International Symposium on Electronic Imaging 2018 Visual Information Processing and Communication IX
4. Kalyan Goswami, Byung-Gyu Kim, Dongsan Jun, Soon-Heung Jung, and Jin Soo Choi ," Early Coding Unit–Splitting Termination Algorithm for High Efficiency Video Coding (HEVC)" , ETRI Journal, Volume 36, Number 3, June 2014J
5. Buddhiprabha Erabadda, Thanuja Mallikarachchi, Gosala Kulupana and Anil Fernand," Machine Learning Approaches for Intra-Prediction in HEVC", 2018 IEEE 7th Global Conference on Consumer Electronics (GCCE 2018)
6. Bochuan Du, Wan-Chi Siu and Xuefei Yang," Fast CU Partition Strategy for HEVC Intra-Frame Coding Using Learning Approach via Random Forests", Proceedings of APSIPA Annual Summit and Conference 2015
7. Mateus Grellert, Bruno Zatt,, Sergio Bampi‡, Luis A. da Silva Cruz§,"Fast Coding Unit Partition Decision for HEVC Using Support Vector Machines", IEEE Transactions on Circuits and Systems for Video Technology
8. Linwei Zhu, Yun Zhang, Zhaoqing Pan, Ran Wang, Member, IEEE, Sam Kwong, Zongju Peng," Binary and Multi-Class Learning Based Low Complexity Optimization for HEVC Encodin", IEEE Transactions On Broadcasting , Accepted for publication
9. Xin Lu∗, Nan Xiao∗, Graham Martin†, Yue Hu ∗ and Xuesong Jin, "Fast Intra Coding Implementation for High Efficiency Video Coding (HEVC)",2017 Data Compression Conference
10. Yih-Chuan Lin, Jian-Cheng Lai," Feature-based fast coding unit partition algorithm for high efficiency video coding", Journal of Applied Research and Technology 13 (2015) 205-219
11. Bouke Timbermont ,"Fast Mode encoding with HEVC: reducing complexity by predicting Coding Unit structure",2013
12. Chan-seobPark,Gwang-SooHong,andByung-GyuKim, " Novel Intermode Prediction Algorithm for High Efficiency Video Coding Encoder", Advances in Multimedia Volume 2014, Article ID 196035, 8 pages http://dx.doi.org/10.1155/2014/196035
13. Sangsoo Ahn, Bumshik Lee and Munchurl Kim, "A Novel Fast CU Encoding Scheme based on Spatio-Temporal Encoding Parameters for HEVC Inter Coding", DOI 10.1109/TCSVT.2014.2360031, IEEE Transactions on Circuits and Systems for Video Technology
14. Yun Zhang, Sam Kwong, Xu Wang, Hui Yuan, Zhaoqing Pan, Long Xu ," Machine Learning-Based Coding Unit Depth Decisions for Flexible Complexity Allocation in High Efficiency Video Coding", Ieee Transactions On Image Processing, Vol. 24, No. 7, July 2015
15. Gary J.Sullivan,Jens-Rainer Ohm, Woo-Jin Han, "Overview of High Efficiency Video Coding (HEVC) Standard", IEEE Transactions on Circuits and Systems for Video Technology,Vol:22, Issue;12,Dec2012
16. Frank Bossen, Benjamin Bross,Karsten Suhring, David Flynn," HEVC Complexity and Implementation Analysis", IEEE Transactions on Circuits and Systems for VideoTechnolgy,Vol.22,Issue12,2012
17. Vivienne Sze, Madhukar Budagavi, Gary Sullivan," High Efficiency Video Coding (HEVC)Algorithms and Architectures", Springer Publication
18. Iain E.G.Richardson," H.264 and MPEG-4 Video Compression Video Coding for nest generation Multimedia", Wiley Publication

AUTHORS PROFILE

Mrs.Chhaya S Pawar received B.E and M.E degree from University of Mumbai in 2006 and 2013 respectively. She has published 28 papers in various national and international journals and conferences and is member of Computer Society of India and International Association of Engineers. She is currently pursuing Ph.D degree from University of Mumbai. Her research interest includes Artificial Intelligence, Video Coding.

Dr.Sudhir Deoraoji Sawarkar received B.E and M.E in Electronics Engineering and Ph.D degree in Computer Engineering. He has more than 29 years of Experience. He has published more than 182 research papers in national and international journals and conferences and is member of Computer Society of India and ISTE. His Research interests include Artificial Intelligence, Neural Networks, Image Processing.