IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 5, May 2022
Impact Factor: 6.252

'Vedic' Sanskrit Language Character Recognition from Images using CNN and OCR

Dr. Sunil L. Bangare[1], Ketan S Gore[2], Ganesh S. Waghmare[3], Bhagyashri Bhoi[4], Mallika Marndi[5]
Associate Professor, Department of Information Technology[1]
UG Scholar, Department of Information Technology[2,3,4,5]
Sinhgad Academy of Engineering, Pune, Maharashtra, India

Abstract: *Many scholars have recently been interested in deep learning and character recognition. Deep neural networks exhibit cutting-edge performance in many classification and identification issues. The Optical Character Recognition (OCR) algorithm takes an optical picture of a character as input and provides the corresponding character with its current meaning and execution time as output. It has several uses, including traffic surveillance, robotics, and the digitalization of printed documents. Convolutional Neural Network (CNN), a prominent deep neural network design, may be used to construct OCR. The standard CNN classifiers are capable of learning the significant 2D characteristics contained in pictures and classifying them using the soft-max layer. The CNN is used to extract features. Several common CNN classifiers were investigated in order to discover optimal CNN for extracting features that may be utilised in combination with ECOC classifier for accurate recognition of handwritten or any character in Sanskrit. The given handwritten character image dataset is used to train and evaluate the CNN-ECOC. The simulation results reveal that CNN provides greater accuracy and somewhat different meaning than the classic CNN classifier.*

Keywords: Character recognition; Classification; CNN; Deep learning; OCR; SVM.

I. INTRODUCTION

OCR is the technique of matching optical patterns in a digital picture to their corresponding characters. Character recognition is accomplished through crucial procedures such as feature extraction and categorization [1]. One of the most successful applications in the field of object recognition is the OCR system, which replicates the human capacity to identify printed forms of text. OCR applications include identifying a vehicle's registration number from an image of the licence plate, which aids in traffic control [2], converting printed academic records into text for storage in an electronic database, decoding ancient scripture, and automatic data entry via optical scanning of cards, bank checks, and so on. OCR technologies reduce typing errors and save time.Handwritten character identification is difficult, and academics have been experimenting with various algorithms over the last few decades. Deep neural networks have recently piqued the interest of many academics because to its undeniable capacity to solve computer vision problems such as object identification, classification, recognition [4], and so on. CNN is a common form of deep neural network that can learn and extract characteristics from 2D pictures. The CNN classifier is capable of recognising characters in a picture. Traditional CNN classifier design comprises of convolutional layers for feature extraction and fully connected layers followed by a soft-max layer for classification. CNN is a powerful feature extractor [5].

In this article, CNN, a hybridization of CNN architecture and ECOC classifier, is given. The CNN is used to extract and recognise features. In CNN-ECOC, the soft-max layer of classic CNN is replaced by the ECOC classifier. ECOC primarily translates multi-class classification problems into grey and binary formats in order to interpret input to the computer or our system problem using various coding schemes and a linear learner such as Support Vector Machine (SVM). The SVM transforms the inputs into a high-dimensional space where the distinctions between classes may be seen. The SVM can automatically prevent overfitting and has a high prediction accuracy [6]. The SVM is also more generalizable than neural networks [7].

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 5, May 2022
Impact Factor: 6.252

Convolutional Neural Networks

CNNs are constructed from a large number of linked neurons with learnable weights and biases. CNN's neurons are structured as layers in its architecture. It is made up of an input layer, many hidden layers, and an output layer. When a network contains a significant number of hidden layers, it is referred to as a deep neural network. CNN's hidden layers link to a tiny part (receptive field) of the input space formed by the preceding layer rather than connecting to all, as in fully connected networks such as Multi Layered Perceptron (MLP) networks.Rectified Linear Unit (ReLU) is a nonlinear activation function. It substitutes negative values with zero, which helps accelerate learning. The activation function is applied to the output of each convolution layer. The pooling layer minimises the spatial size of each feature map, resulting in less processing in the network. Pooling also employs a sliding window that glides over the feature map in stride to turn it into representative values. The terms minimum pooling, average pooling, and maximum pooling are often used. Every neuron in the layer is fully linked to every neuron in the previous layer.

AlexNet

AlexNet has 25 layers. It has five convolution layers, followed by a ReLU layer and a maximum pooling layer. The cross-channel normalising layer is added after the first two convolution layers. Cross channel normalisation is equivalent to channel-wise normalisation. It substitutes a normalised value acquired from neighbouring cells for each element. It also contains three completely connected layers, which are followed by the ReLU Layer, and the first two fully connected levels are also followed by dropout layers. The output of the last fully connected layer is sent into softmax, which generates a probability distribution with 1000 classes. AlexNet was created to recognise objects in ImageNet [11].

ZfNet

ZfNet's design is similar to AlexNet's, however the filter size and number of filters utilised varies. In the first convolution layer, the AlexNet employs a filter of size 11 11 with a stride of 4, whereas the ZfNet employs a filter of size 7 7 with a stride of 4. In the third, fourth, and fifth convolution layers, the ZfNet employs 512,1024, and 512 filters, respectively, whereas the AlexNet employs 384,384, and 256 filters. The ZfNet was also built for object recognition in ImageNet [11].

LeNet

As an input, LeNet accepts a grayscale picture of size 32 32. The design is made up of two sets of convolutional and average pooling layers, then a flattening convolutional layer, two fully connected layers, and a softmax layer for classification. The LeNet is a pioneering study in utilising CNN to recognise handwritten numbers from photographs. Many studies have presented various CNN architectures for identifying handwritten characters, with great accuracy.

Working of CNN with OCR

Convolutional Neural Network (CNN) is a Deep Learning method that can take in an input image, give importance (learnable weights and biases) to distinct aspects/objects in the image, and distinguish one from the other. After executing the algorithm, create a model and proceed to the testing step to predict the result. Figure 1 depicts the stages involved in character recognition. The input picture is fed into the CNN, which extracts the features, tests all essential cases and characteristics investigated by the extraction, and produces the matching character as an output after testing.

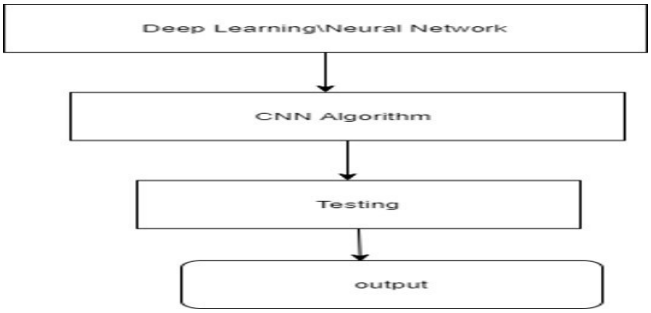


Figure 1: CNN-OCR Working

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 5, May 2022
Impact Factor: 6.252

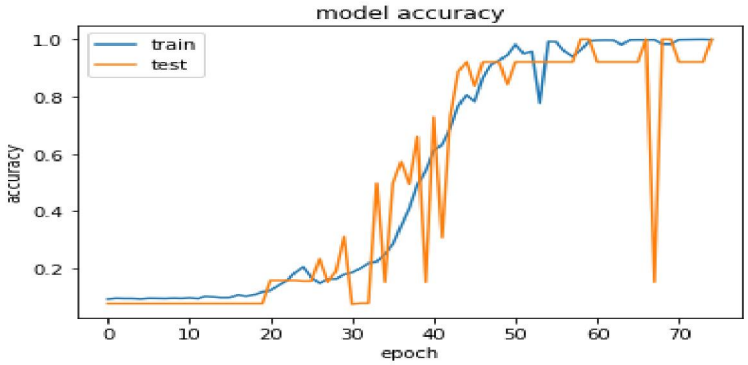
II. IMPLEMENTATION AND RESULT

For feature extraction, four distinct CNN architectures have been constructed. The features are extracted from the output of the CNN's final fully linked layer. The CNN and ECOC were trained using the NIST handwritten character dataset [17]. Ciresan [7] presented handwritten character recognition using CNN and ECOC in 2011 and Fanany in 2017 [18]. Training and testing on the NIST dataset yielded accuracy of 88 percent and 93 percent, respectively. The goal of this work is to improve the accuracy of the CNN character recognition system by using the ECOC classifier. The dataset is divided into 26 folders, each containing 2473 (1483 training images and 990 testing images) different upper-case English alphabet handwritten or image character images.

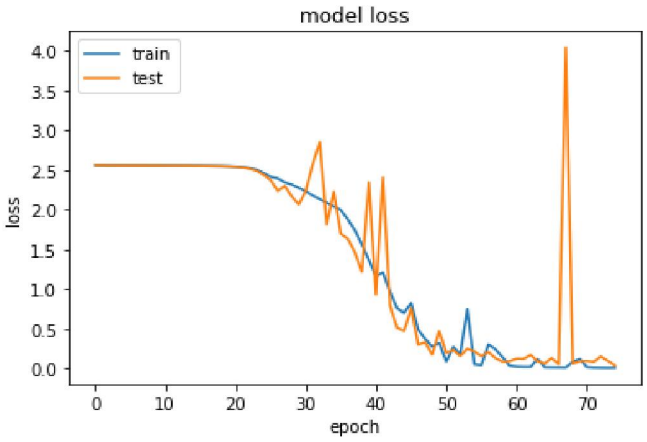
LeNet of Type 1 with ECOC Approach

A grayscale picture of size 32 32 is fed into the LeNet of type one. The gradient descent back propagation approach was used to train it, and the mini-batch size and maximum epoch were set to 64 and 20, respectively. The training accuracy rate was 74.63 percent, while the testing accuracy rate was 73.78 percent. The characteristics are then retrieved from the trained LeNet's final fully connected layer and put into an ECOC classifier.The mini-batch size and max epoch are set to 64 and 20, respectively, and training and testing accuracy rates of 89.00 percent and 85.86 percent have been attained. Replacing the softmax layer with ECOC improves classification accuracy in the majority of networks. If the features are retrieved using the AlexNet architecture, the ECOC classifier has the best testing accuracy. S. L. Bangare et al. [20-26] studied brain tumour detection. N. Shelke et al [27] proposed the LRA-DNN approach. Suneet Gupta and colleagues [28] worked on an end-user system. Gururaj Awate et al. [29] investigated Alzheimer's disease. P. S. Bangare and colleagues [30-32] worked on object detection.Kalpana Thakare et al [33-38] have worked on various machine learning algorithms. M. L. Bangare et al. [39-40] worked on the cloud platform. Rajesaheb R. Kadam et al [41] and Sachindra K. Chavan et al. [42] have discussed security issues with cloud.

Accuracy Graph:



Model Loss Ratio:



IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 5, May 2022
Impact Factor: 6.252

Here in below added images of some current working state of this system are as;

III. RESULTS

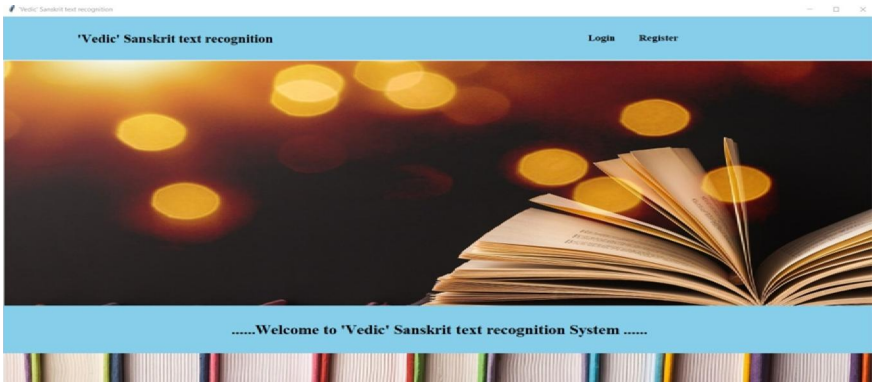


Fig. 2: Main File

Fig 3: Registration File

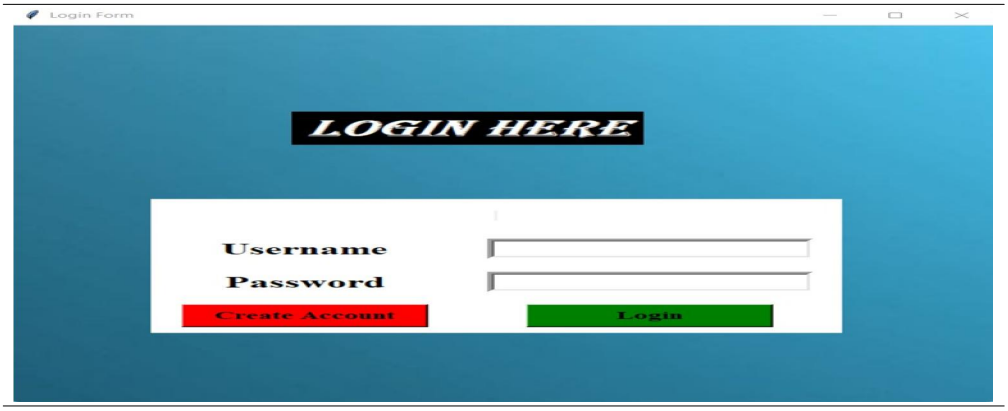


Fig 4: Login File

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 5, May 2022
Impact Factor: 6.252

Fig 5: Master

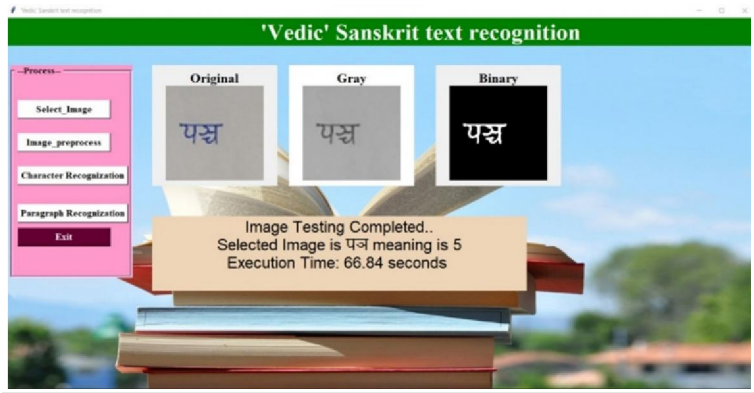


Fig 6: Check File

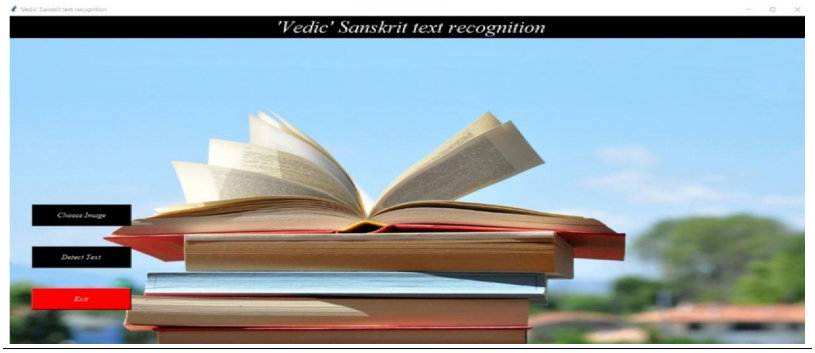


Fig 7: Check File

IV. CONCLUSION

Handwritten character recognition methods employing CNN-ECOC, a hybrid of CNN and ECOC classifiers, are given in this article. The CNN is utilised for feature extraction, while the ECOC is used for character recognition. Three common CNN architectures, notably LeNet, AlexNet, and ZfNet, have been investigated in order to locate a suitable feature extractor. Based on the simulation results, it was discovered that LeNet had a poor accuracy rate. As a result, it has been changed by adding a dropout layer and a ReLu layer after the first completely linked layer, resulting in a greater accuracy rate. It has also been discovered that ECOC classifiers outperform CNN softmax classifiers in terms of accuracy.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 5, May 2022
Impact Factor: 6.252

REFERENCES

[1]Chaudhuri, Arindam and Mandaviya, Krupa and Badelia, Pratixa and Ghosh, Soumya K and others. (2017) "Optical Character Recognition System. In Optical Character Recognition Systems for Different Languages with Soft ComputingSpringer: 941.

[2]Li, Haixiang and Yang, Ran and Chen, Xiaohui. (2017) "License plate detection using convolutional neural network. 3rd IEEE International Conference on Computer and Communications (ICCC),IEEE:17361740.

[3]Rajavelu, A and Musavi, Mohamad T and Shirvaikar, Mukul Vassant. (1989) " A neural network approach to character recognition.Neural Network 5,Elsevier (2): 387393.

[4]Bai,Jinfeng and Chen, Zhineng and Feng, Bailan and Xu, Bo.(2014) "Image character recognition using deep convolutional neural network learned from different languages. IEEE International Conference on Image Processing (ICIP):25602564.

[5]Maitra, Durjoy Sen and Bhattacharya, Ujjwal and Parui, Swapan K. (2015) "CNN based common approach to handwritten character recognition of multiple scripts.13th International Conference on Document Analysis and Recognition (ICDAR),IEEE:10211025.

[6]Jakkula,Vikramaditya. (2006)"Tutorial on support vector machine (svm). School of EECS, Washington State University 37.

[7]Ciresan,Dan Claudiu and Meier,Ueli and Gambardella,Luca Maria and Schmidhuber,Jurgen. (2011)"Convolutional neural network committees for handwritten character classification. International Conference on Document Analysis and RecognitionIEEE :11351139.

[8]Krizhevsky, Alex and Sutskever, Ilya and Hinton, Geoffrey E. (2012) "Imagenet classification with deep convolutional neural net- works.Advances in neural information processing systems:10971105.

[9]Zeiler, Matthew D and Fergus,Rob. (2014) "Visualizing and understanding convolutional networks. European conference on computerSpringer, vision:818833.

[10]LeCun, Yann and Bottou, Leon and Bengio, Yoshua and Haffner, Patrick and others. [11](1998)"Gradient-based learning applied to document recognition. Proceedings of the IEEE,Taipei, Taiwan, 86 (11): 22782324.

[12]Guyon, Isabelle and Schomaker, Lambert and Plamondon, Rejean and Liberman, Mark and Janet, Stan. (1994) "UNIPEN project of on-line data exchange and recognizer benchmarks." Proceedings of the 12th IAPR International Conference on Pattern RecognitionIEEE,(2):2933.

[13]Yuan, Aiquan and Bai, Gang and Jiao, Lijing and Liu, Yajie. (2012) "Onine handwritten English character recognition based on convolutional neural network. 10th IAPR International Workshop on Document Analysis SystemsIEEE: 125129.

[14]Rahman, Md Mahbubar and Akhand, MAH and Islam, Shahidul and Shill, Pintu Chandra and Rahman, MH and others. (2015) "Bangla hand- written character recognition using convolutional neural network. International Journal of Image, Graphics and Signal Processing(IJIGSP),7 (8): 4249.

[15]Deng, Huiqun and Stathopoulos, George and Suen, Ching Y. (2009) "Errorcorrecting output coding for the convolutional neural network for optical character recognition. 10th International Conference on Document Analysis and Recognition,IEEE: 581585.

[16]Deng, Huiqun and Stathopoulos, George and Suen, Ching Y. (2010) "Applying error-correcting output coding to enhance convolutional neural network for target detection and pattern recognition.20th International Conference on Pattern Recognition,IEEE :42914294.

[17]Dietterich, Thomas G and Bakiri, Ghulum. (1994) "Solving multiclass learning problems via error-correcting output codes. Journal of artificial intelligence research (2):263286.

[18]Grother,Parick J and Hanaoka,Kayee K. (2016) "NIST special database 19 handprinted forms and characters database. National Institute of Standards and Technology.

[19]Fanany,Mohamad Ivan and others. (2017) "Handwriting recognition on form document using convolutional neural network and support vector machines (CNN-SVM). 5th International Conference on Information and Communication Technology (ICoIC7),IEEE:16.

[20] S. L. Bangare, G. Pradeepini, S. T. Patil, "Implementation for brain tumor detection and three dimensional visualization model development for reconstruction", ARPN Journal of Engineering and Applied Sciences (ARPN JEAS),

Vol.13, Issue.2, ISSN 1819-6608, pp.467-473. 20/1/2018 http://www.arpnjournals.org/jeas/research_papers/rp_2018/jeas_0118_6691.pdf

[21] S. L. Bangare, S. T. Patil et al, "Reviewing Otsu's Method for Image Thresholding." International Journal of Applied Engineering Research, ISSN 0973-4562, Volume 10, Number 9 (2015) pp. 21777-21783, © Research India Publications https://dx.doi.org/10.37622/IJAER/10.9.2015.21777-21783

[22] S. L. Bangare, G. Pradeepini, S. T. Patil, "Regenerative pixel mode and tumor locus algorithm development for brain tumor analysis: a new computational technique for precise medical imaging", International Journal of Biomedical Engineering and Technology, Inderscience, 2018, Vol.27 No.1/2. https://www.inderscienceonline.com/doi/pdf/10.1504/IJBET.2018.093087

[23] S. L. Bangare, A. R. Khare, P. S. Bangare, "Quality measurement of modularized object oriented software using metrics", ICWET '11: Proceedings of the International Conference & Workshop on Emerging Trends in Technology, February 2011, pp. 771–774. https://doi.org/10.1145/1980022.1980190.1.

[24] S. L. Bangare, G. Pradeepini and S. T. Patil, "Brain tumor classification using mixed method approach," 2017 International Conference on Information Communication and Embedded Systems (ICICES), 2017, pp. 1-4, doi: 10.1109/ICICES.2017.8070748.

[25] S. L. Bangare, S. Prakash, K. Gulati, B. Veeru, G. Dhiman and S. Jaiswal, "The Architecture, Classification, and Unsolved Research Issues of Big Data extraction as well as decomposing the Internet of Vehicles (IoV)," 2021 6th International Conference on Signal Processing, Computing and Control (ISPCC), 2021, pp. 566-571, doi: 10.1109/ISPCC53510.2021.9609451.

[26] S. L. Bangare, G. Pradeepini, S. T. Patil et al, "Neuroendoscopy Adapter Module Development for Better Brain Tumor Image Visualization", International Journal of Electrical and Computer Engineering (IJECE) Vol. 7, No. 6, December 2017, pp. 3643~3654. http://ijece.iaescore.com/index.php/IJECE/article/view/8733/7392

[27] N. Shelke, S. Chaudhury, S. Chakrabarti, S. L. Bangare et al. "An efficient way of text-based emotion analysis from social media using LRA-DNN", Neuroscience Informatics, Volume 2, Issue 3, September 2022, 100048, ISSN 2772-5286, https://doi.org/10.1016/j.neuri.2022.100048 .

[28] Suneet Gupta, Sumit Kumar, Sunil L. Bangare, Shibili Nuhmani, Arnold C. Alguno, Issah Abubakari Samori, "Homogeneous Decision Community Extraction Based on End-User Mental Behavior on Social Media", Computational Intelligence and Neuroscience, vol. 2022, Article ID 3490860, 9 pages, 2022. https://doi.org/10.1155/2022/3490860.

[29] Gururaj Awate, S. L. Bangare, G. Pradeepini and S. T. Patil, "Detection of Alzheimers Disease from MRI using Convolutional Neural Network with Tensorflow",arXiv, https://doi.org/10.48550/arXiv.1806.10170

[30] P. S. Bangare, S. L. Bangare, R. U. Yawle and S. T. Patil, "Detection of human feature in abandoned object with modern security alert system using Android Application," 2017 International Conference on Emerging Trends & Innovation in ICT (ICEI), 2017, pp. 139-144, doi: 10.1109/ETIICT.2017.7977025.

[31] P. S. Bangare and S. L. Bangare. "The Campus Navigator: An Android Mobile Application." International Journal of Advanced Research in Computer and Communication Engineering 3, no. 3 (2014): 5715-5717.

[32] P. S. Bangare, N. J. Uke, and S. L. Bangare, "An approach for detecting abandoned object from real time video." International Journal of Engineering Research and Applications (IJERA) 2.3 (2012): 2646-2649.

[33] Kalpana S. Thakare, Viraj Varale, "Prediction of Heart Disease using Machine Learning Algorithm", Bioscience Biotechnology Research Communications (Special issue) Volume 13, Issue 12, 2020 (Dec 2020 issue).

[34] Kalpana S. Thakare, A. M. Rajurkar, "Shot Boundary Detection of MPEG Video using Biorthogonal Wavelet Transform", International Journal of Pure and Applied Mathematics, Volume 118, No. 7, pp. 405-413, ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version), url: http://www.ijpam.eu

[35] Kalpana S. Thakare, A. M. Rajurkar, R. R. Manthalkar, "Video Partitioning and Secured Key frame Extraction of MPEG Video", Proceedia Computer Science Journal, Volume 78, pp 790-798, Elsevier, 2016. Scopus DOI: htpp://10.1016/j.procs.2016.02.058, www.sciencedirect.com/science/article/pii/S1877050916000600

[36] Kalpana S. Thakare, A. M. Rajurkar and R. R. Manthalkar, "Content based Video Retrieval using Latent Semantic Indexing and Color, Motion and Edge Features", International Journal of Computer Applications 54(12):42-48, September 2012, Published by Foundation of Computer Science, New York, USA. DOI: 10.5120/8621-2486

[37] Kalpana S. Thakare, Archana M. Rajurkar, R. R. Manthalkar, "A Comprehensive System Based on Spatiotemporal Features Such as motion, Quantized Color and Edge Features", International Journal of Wireless and Microwave Technologies (IJWMT) ISSN 1449 (Print), ISSN: 2076-9539 (Online), Vol.1, No.3, June. 2011, DOI: 10.5815 /ijwmt

[38] Kalpana S. Thakare, Archana M. Rajurkar, Dr. R. R. Manthalkar, "An effective CBVR system based on Motion, Quantized color and edge density features", International Journal of Computer Science & Information Technology (IJCSIT), ISSN 0975 – 3826, Vol 3, No 2, April 2011 DOI: 10.5121/ijcsit.2011.3206 78.

[39] M. L. Bangare, "Attribute Based Encryption And Data Integrity For Attack on Cloud Storage", Journal of Analysis and Computation (JAC), (An International Peer Reviewed Journal), www.ijaconline.com, ISSN 0973-2861, ICASETMP-2019, pp.1-4. http://www.ijaconline.com/wp-content/uploads/2019/07/ICASETMP67.pdf

[40] M. L. Bangare, Sarang A. Joshi, "Kernel interpolation-based technique for privacy protection of pluggable data in cloud computing", International Journal of Cloud Computing, Volume 9, Issue 2-3, pp.355-374, Publisher Inderscience Publishers (IEL).

[41] Rajesaheb R. Kadam and Manoj L. Bangare, "A survey on security issues and solutions in live virtual machine migration", International Journal of Advance Foundation and Research in Computer (IJAFRC), (December, 2012). ISSN (2014), pp.2348-4853.

[42] Sachindra K. Chavan, Manoj L. Bangare, "Secure Data Storage in Cloud Service using RC5 Algorithm", International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878, Volume-2, Issue-5 November 2013, pp.139-144.