Insights on Video Compression Strategies using Machine Learning

Veena S.K, Mahesh K Rao

Abstract— ***With the rising advancement of the multimedia technology, video compression is becoming a challenging problem. Although, there is availability of various standard compression algorithms, yet robust compression performance is yet to be seen in existing compression techniques. This paper also highlights that machine learning plays a significant contributory role in improving the performance of the video compression. Therefore, this manuscript offers a technical insight about the performance of existing video compression technique using machine learning approach. The contribution of this paper is its findings which states that machine learning approach do have significant advantage but the advantageous features are limited by the inherent and unsolved research problem. The core findings of this paper are basically to highlight the strength and limitations of existing methods as well as to highlight the research gap in terms of open-end research problems which requires immediate attention.***

Keywords: Video Compression, High Efficiency Video Coding, Machine Learning, Encoder, Decoder.

I. INTRODUCTION

The boon of technological advancement is more realized in the area of multimedia which penetrates the global market in a comprehensive manner [1]. Apart from this, a multimedia file is one of the most utilized file systems concerning storage, file sharing, and files editing [2]. This paper speaks about the significance and research trends of video compression which is one of the most challenging and desired requirements in the area of multimedia technologies [3]. There is no denying of the fact that there is enormous research work done towards image compression [4] [5], audio compression [6], and video compression [7]. However, study towards video compression is comparably lesser than the study on image and audio compression. Video compression is one of the essential operations which to be performed for many applications that demand to capture the live feeds 24/7. Modern robotic guided operation in an advanced healthcare system, the live video feeds are required to be stored and transmitted too [8]-[10]. Therefore, video compression positively effects on storage optimization.

At present, there are various compression standards of video which operates on the large scale.

Although, there are various video/multimedia compression scheme till date, but they suffer from following problems viz. i) degradation of signal quality even after using lossless compression, ii) involvement of a considerable number of encoding steps thereby consuming time of encoding, iii) dependencies of multiple non-linear parameters which induces burden on the computation process [11]. Therefore, to perform an effective video compression, it is essential to find the reasons and all the dependable variables that can directly or indirectly assist in superior video compression performance. A superior form of video compression system can be only said when the compressed file size is optimally small without losing signal quality after decompression is carried out. Hence, this problem could be effectively solved using machine learning because machine learning is capable of performing regression, classification, and clustering [12]. These set of operation are highly required in order to find out all the information about the dependable parameters responsible for upgrading compression performance. At present, the potential of machine learning is realized by the various reputed organization to solve more significant end problems [13]-[19]. This is an evidential factor to state that applicability of machine learning is highly more in the area of video compression too. Basically, machine learning offers improved analytical capabilities on complex data that can perform automation of the analytical model. This mechanism is capable of evaluating the algorithms consistently performs learning of the data. These characteristics allow the machine learning to discover all the hidden traits of the data and thereby intelligence is developed on that basis. Some of the potential beneficial features of machine learning schemes are i) capability to process a larger quantity of complex data, ii) extremely faster processing and dynamic prediction capability, iii) exponentially higher applicability. However, they also have many inherent limitations too.

Therefore, this paper presents a discussion of existing machine learning approaches to offer insight into the effectiveness as well as a real picture of the implications of the machine learning approach on video compression. The organization of the paper is as follows- Section II discusses the critical information about video compression followed by a briefing of how machine learning could improve the video compression performance in Section III.

Revised Manuscript Received on October 15, 2019.

Veena S.K, Assistant Professor & Research Scholar, Department of ECE, MITM, Mysuru, Karnataka, India,

Email: veenask_ece@mitmysore.in

****Dr. Mahesh K Rao,** Professor & HOD, Department of ECE, MITM, Mysuru, Karnataka, India,

Email: maheshkrao_ece@mitmysore.in

This section also discusses various standard approaches of video compression using the standard concept of theoretical machine learning and also offers a short discussion of the strength and limitations of existing research contribution. Section-IV briefs an open-end research problem followed by conclusion in Section V.

II. ESSENTIALS OF VIDEO COMPRESSION

A video compression scheme is concerned about the size factor of the video file such that it can be minimized while storing or sending to a different terminal. This is one of the standard procedures used for performing encoding operation over the digital multimedia system. By controlling the size of the video data, there are various benefits, e.g. optimizing the channel capacity during data transmission, fitting the encoded file to the limited storage system, and faster data transmission over peak traffic condition. However, the standard process of video compression does not consider all the factors that are required essential for viewing the output video. At present, there are various forms of the codecs responsible for performing compression operation with a balance towards end quality of the video. Usage of codec also ensures that there is an efficient use of the hardware resources while performing compression operation. There are different levels of compression operation that can be carried out in one codec of compression [20].

According to all recent studies, the process of video compression can be classified in two forms, i.e. i) intra-frame compression (e.g., Motion-JPEG) and ii) inter-frame compression (MPEG-1, MPEG-2, MPEG-4, H.264). In the intra-frame approach, only the current frame of video is subjected to compression while inter-frame approach involves many frames for compression [23] [24]. Many numbers of frames could be either frame that is before or after the current frame of the given video. The new format of the video compression is H.265 (also known as High-Efficiency Video Coding) [25] and VP8 [26]. It is believed that H.265 has the capability to compression 25% of the file size with a reduction in computational overhead of 50%. However, it is still under the process of development. VP8 is the new codec owned by Google which claims of highly reduced data usage compared to H.264 and also optimize channel capacity with approximated 40% of performance enhancement. However, it is yet to be standardized for commercial utilization globally.

Table.1 Summary of Strength and Limitations of Existing Video Compression Standards

Compression Standard	Compression Factor	Strength	Limitations
Motion-Joint Photographic expert group (M-JPEG)	1-20	-distinct and flawless image when frame-rates are less in contrast to MPEG-4 -highly reduced utilization of CPU -insensitive to complexity factor motion	Less efficient as compared to H.264 and MPEG-4
Moving Picture Expert Group (MPEG-4)	1-50	-highly support wide variety of devices (mobile/digital) -Better quality of broadcasting data in television as well as streaming video	-non-effective compression -maximized utilization of CPU
H.264	1-100	-highly efficient for video with less motion	-non-effective compression -highly maximized utilization of CPU

However, there are various interchangeable terms like video codec, video compression, and video container when dealing with this topic. Basically, it is required to understand that a video file consists of multiple components e.g. video stream, audio stream, and meta-data, which are first subjected to compression followed by encapsulation operation. This leads to generation of coded video file which is compressed. The currently used video compression schemes are H.263, H.264, MPEG (2/4), Sorenson spark, VP6, etc. The currently used audio compression schemes are MP3, AAC, Vorbis, etc while compression used for meta-data are RDF, XML, XMP, etc. Basically, a video codec consists of instructions that are responsible for determining the process deployed for compression as well as reverse steps in de-compression operation. Existing codec are of two types viz. lossy and lossless. Apart from all these, there are two more non-conventional video codec called as DV [21] and Huffyuv [22]. DV is basically a lossy compression technique and it is much superior to Motion JPEG in all sense with supportability on cross-platform device. However, it has some inherent problems too with respect to resolution, frame-rate, and data rate that are highly restrictive. Huffyuv is basically lossless compression standard and it offers a higher degree of retention of signal quality. An interesting fact about this approach is that irrespective of multiple times of compression operation using this codec, the size of the video data is not going to change. However, the performance of decompression is very different than compression operation. The time taken during decompression operation in this code is quite higher than that of compression operation. It also suffers from maximized dependency of the storage space. This will mean that a Huffyuv is better application when a compression and decompression is carried out on different device with different application in order to retain nearly similar processing time. Hence, this is not limitation and rather is a constraint.

Therefore, there are various newly evolved variants of video codec that claims of some advantageous features as well as also found with some limitations too when experimentally analyzed. The next section discusses about the trends of existing literatures.

III. IMPACT OF MACHINE LEARNING IN COMPRESSION

The positive effect of using machine learning in performing compression is not new. In 2006, Sculley [27] has already briefed about the advantages that machine learning could bring an improvement in the performance of compression. Apart from this, machine learning has been increasingly adopted in developing Artificial Intelligence in the area of multimedia [28] [29]. This part of the manuscript establishes a link between the machine learning and video compression and avoids any repetitive discussion of any theoretical facts of them.

A. Standard Scheme of Machine Learning

It is well known that machine learning is essentially used for performing prediction and this predictive operation is something that can be assistive while developing a robust compression system of multimedia. It should be noted that majority of the machine learning tools that carry out compression of video deploy statistical predictive techniques which is slightly different from other mainstream predictive approach. Fig.1 highlights the conventional form of predictive operation that is involved in both machine learning and statistical modeling inclusive of the all the essential factors that are deployed during the training operation. Generally speaking, a good training data as well as an effective mathematical model is essential for developing a robust predictor [30]. This can assist in enhancing the accuracy measures along with reduction in computational complexity. However, this is just a claim offered by theory of machine learning.

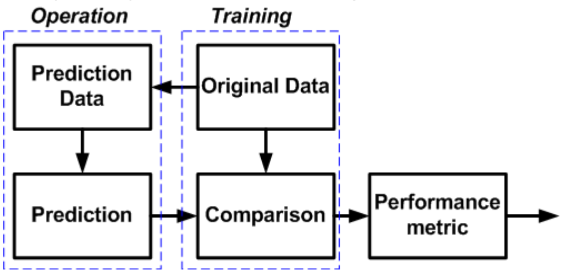


Fig.1 Conventional Predictive Scheme

At present, there are potential technological advancement that is capable of constructing sophisticated predictors for assisting in complex predictive problems by machine learning [31]- [35]. This process also generates a massive training data which could be either directly implemented or slightly customized and then implemented. There is no doubt that if such complex system of machine learning is applied on video compression than some superior and outstanding results can be expected in the compression performance too. Fig. 2 highlights a generalized scheme where machine learning is applied in order to carry out compression. This figure is constructed on the basis of generalization of different implementation concept presented by various researchers till day. This scheme offers an advantage of re-thinking about any possible adding of block of operation or editing an existing block of operation depending upon the requirement of the user. The complete figure has two blocks viz. encoder and decoder block. It can be seen that original video data is subjected to encoder block which after applying encoding compresses the file and complies with the rate-distortion optimization. On the other hand, the decoder is responsible for data recovery on the basis of the prediction block.

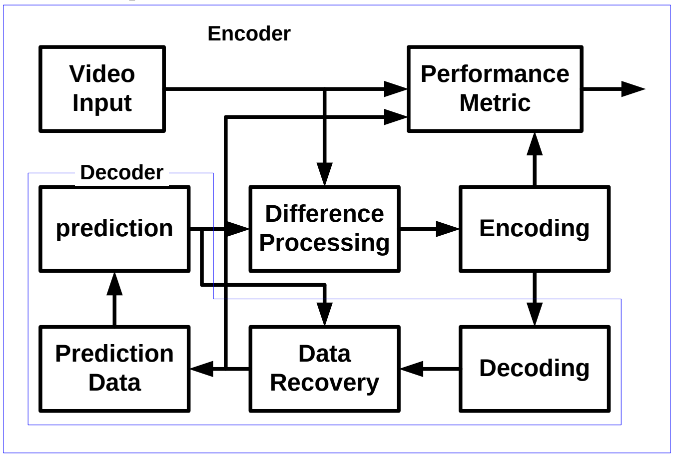
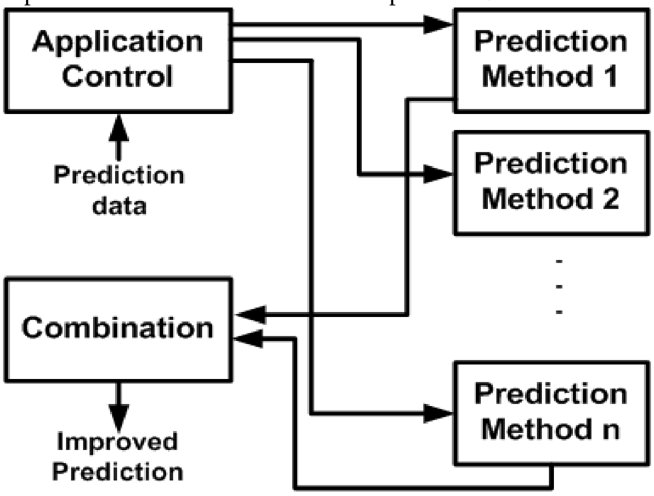


Fig.2 Significance of Prediction in Compression

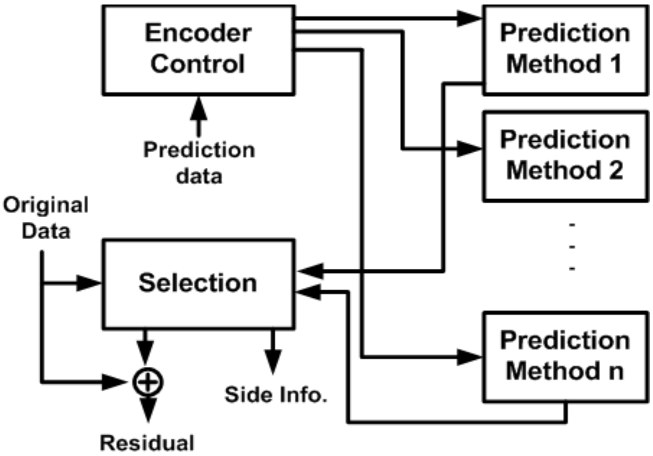
As the amount of the data to be compressed is directly proportional to the mean difference between the actual video data and predicted video data, therefore, such mechanism always consider that better prediction should be employed in order to achieve better compression performance. Therefore, Fig.2 basically depicts the importance of prediction in compression. Once the predictor is identified then it is simple job to use different ranges of machine learning model in order to perform training operations on those predictors. However, the outcomes are not always guaranteed to offer best results. This is a complex problem and one way to solve this problem using machine leaning is to actually use a correct direction of investigation. A closer look into Fig.2 shows that encoder block is always having a direct accessibility to the input video file as well as decoder is not a separate module but is considered as a sub-module of an encoder itself. This is an interesting finding very different from other conventional encoding decoding schemes.

The above stated observations are basically utilized in order to re-consider the effect of predictors as the conventional scheme always permits the encoder system to compare its outcome with the decoder information as well as original video file in order to obtain correct decision. For better understanding, Fig.3 offers more comprehensive insights with respect to generalized prediction (Fig.3(a)) and video compression (Fig.3(b)). This scheme shows a conventional mechanism of utilization of different predictors. Different outcomes of the prediction operation are shown to be integrated in the Fig.3(a) where a specific confidence measure is used to derive diversified weights. Owing to non-availability of original video input file, the system is actually not capable of confirming the best predictor to be considered. It should be known that original video file is only considered when training operation is performed and not after that.

One effective way to solve this problem is to use a combinatorial approach [36] which can enhance the predictive performance. This operation could lead to selection of best predictor which should be considered for further selection operation for obtaining best result. On the other hand, a discrete encoding system is shown in Fig.3(b) where original video source can be accessed which makes the task quite easier to explore the best and precise predictive performance. However, it should be known that accuracy factor shouldn't be the only parameter for best compression performance. The main reason behind this is there is always a demand of additional data bit for the encoder block in order to update the decoder block about the predictor that is adopted. By doing this, the overhead has to be considered as well as rate distortion factor should be well represented in order to find the best predictor.



(a) General Prediction



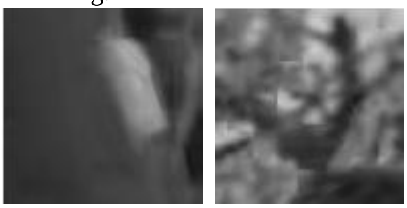
(b) Compression

Fig.3 Operation Flow of Machine Learning in Compression

By this discussion, it is essential to understand that predictors have different impact on different problems and adherence to conventional scheme of machine learning is always good idea. However, the problems associated with using current video codec like HEVC offers different impediment that doesn't allow the conventional scheme to offer claimed benefits. Hence, existing researchers have evolved up with various solution to address this problem.

B. Typical Machine Learning Approach

A conventional approach for multimedia compression calls for downscaling and up-scaling the resolution factor prior to and after compression. This approach is mainly used to smartly optimize the channel capacity. However, this approach is quite designed to meet very specific requirement of particular application. Most recently, the work carried out by Li et al. [37] has addressed this problem using convolution neural network. The main idea of this paper is to showcase that the recent video compression standard of HEVC can be out-performed where the presented technique is about obtaining a high-resolution frame from lower version of it. Although, this idea offers a better alternative of conventional compression approach but they are never attempted on video files. Most recent study towards video compression was carried out by Ryu and Kang [38] where intra-prediction is carried out using machine learning approach. An arbitrary tree is developed where learning operation is carried out following by reducing the size of the candidate modes prior to perform optimization. The operation was carried out over HEVC and it aimed for minimizing the encoding time. In many studies, it has been shown that convolution neural network is well known for its capability to control artifacts caused due to JPEG compression; however, it is less discussed that they are not able to handle much more complex artifacts that appears in new compression standards. This problem is addressed in work of Soh et al. [39] where the conventional convolution neural network is enhanced by considering the temporal correlational factor obtained from the video file. The presented system claims of addressing artifacts based on directional patterns in HEVC (Fig.4(a)) and artifacts of blocking in HEVC (Fig.4(b)). Both the forms of artifacts are quite detrimental for required signal quality while performing decoding.



(a) (b)

Fig.4 Artifacts of HEVC

Study adopting deep neural network was also witnessed in the work of Chen et al. [40] where a convolution neural network is implemented for performing video compression. The technique also applies scalar quantization as well as standard Huffman encoding system in order to obtain feature map. Usage of convolution neural network is also seen in the work of Jiang et al. [41] where a comprehensive learning approach is implemented focusing in compactness and reconstruction phase of it. The machine learning algorithm has been implemented for overcoming the gradient-based problem in quantization step. Although, the study is focused on video stabilization [42] but the idea of implementing the learning scheme just shows that adoption of such learning scheme not only assists in video compression but also in improving its quality.

The study is validated over

standard video compression H.264. Another essential observation is that there is lesser number of studies where compression is actually associated with the networking factors. The work of Cheng et al. [43 has considered this fact where a video encoding operation is carried out on the basis of the coding rate. Therefore, an adaptive process is presented for transmitting video in compressed version. The interesting part of the proposed system is that it assesses the quality of the video on the basis of network parameters captured from standard transfer protocol.It has been also observed that maximum work is carried out considering HEVC as it is the upcoming video codec standard. However, in reality, majority of the devices still uses its prior version i.e. H.264 only. Hence, transcoding is one such mechanism that can transform the video encoding mechanism supporting from older to newer version without using any extra hardware. Most recently, the work carried out by Lin et al. [44] have constructed a classifier using tree concept and Bayesian approach. The prime intention of the study was to forecast the depth of coding unit in HEVC using a unique feature selection method. A unique study towards considering encoding of three-dimensional video is carried out by Zhu et al. [45] where convolutional neural network is implemented for improving the visual quality of the video. The input to the system is a texture video and depth video which is subjected for three-dimensional encoding using HEVC. The encoded information is transmitted to the decoder which generates virtual viewpoints. However, there are various authors who critically commented the better quality of encoding of HEVC codec associated with computational burden. The hypothesis presented by Iranfar et al. [46] claimed of connectivity between the computational burden and thermal factor of the cores. The authors have presented a machine learning-based mechanism where dynamic learning operation is carried out in order to perform selection of an elite encoding system allocated for individual video files. The authors have presented assignment-based strategy, which consider the resolution required as well as an effective migration strategy. This work has proven that enhancing the learning method has significant effect on the encoding mechanism. Nearly similar forms of methodology were also adopted by Gao e.t. al. [47] where emphasis was given to achieve benchmark outcome considering rate-distortion theory over quantization parameter. The problem associated with the HEVC operation with respect to computational complexity has been addressed in the work of Xu et al. [48]. According to the author, deep learning approach offers better prediction scheme with lesser burden on encoding. The scheme addresses the complexity associated with both inter and intra mode prediction in HEVC where the backbone technique used is convolution neural network and memory-based network.

In the existing system, there is also a report of complexity of HEVC due to coding tree structure as well as deployment of various complicated tools for coding. This issue has been addressed by Zhu et al. [49] [50] where a fuzzy logic has been used along with support vector machine. The fuzzy logic is used for developing decision process of coding unit and complexity of optimization process of rate distortion is addressed using support vector machine. The technique also addresses the uncertainty problem with the aid of section of set of features on the basis of inappropriate classification with respect to cost / risk factor. The outcome of the study is claimed for approximately 50% of computational complexity reduction. Similar problems of complexity associated with the HEVC was addressed by Wang et al. [51] where a quality of a video is subjected to improvement in the decoder part unlike other existing system. The authors have also used convolution neural network for eliminating the possible artifact presents. The system also exploits all sorts of under-utilized information residing in the bit-streams for further improving the performance of compression. Hence, the study is more focused on achieving better decoding performance. Hence, neural network and its different advanced variants are used more in solving the complexities associated with the HEVC performance. However, different from such mainstream study, the authors Gao et al. [52] have used game theory integrated with the machine learning. The authors have used support vector machine in order to enhance the accuracy factor related to the classification associated with the rate distortion model. The presented approach uses Nash Equilibrium for optimizing the process of allocating the bits on the stream. The adjustment is carried out over the inter and intra frames of the quantization parameters for better optimization performance. Liu et al. [53] have presented a study where an algorithm is formulated for adaptive decision making associated with the intra-prediction mode. The paper basically addresses the problem associated with the iterative operation of HEVC leading to computational complexity. The study, therefore, introduces a unique coding unit design which works fast and is highly adaptive using machine learning mechanism. The complexity factor is controlled by extracting the image features followed by classification operation on supportive vector machine. Study towards fast mode prediction in intra frame was carried out by Duanmu et al. [54] where the authors have implemented machine learning for developing decision framework and fast mode. According to this process, the coding unit is divided into two different blocks on the basis of screen content and natural image. An explicit coding scheme is introduced for these two images and hence a classified mechanism of distinct coding is introduced. Therefore, the authors have emphasized mainly on decision towards partition and fast mode unlike existing system which is more inclined towards fast mode only. Further, adoption of machine learning is considered where the input is divided into non-overlapping unit of coding tree and coding unit. A comprehensive search process is implemented for determining encoding parameters.

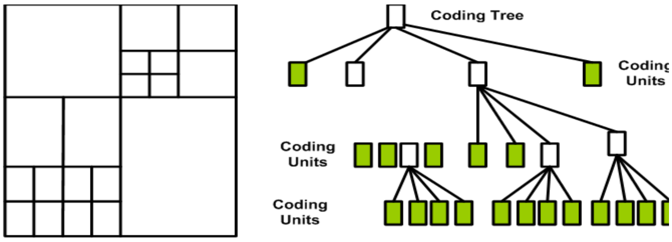


Fig.5 Structure for partitioning screen contents [54]

Study towards adaptive encoding system is carried out by Van et al. [55] where an integration of the distributed as well as predictive scheme of coding is introduced. The authors have presented the solution towards the complexity of the video coding system using two discrete approaches. The first approach is related to the machine learning while the second approach is related to the correlation-based approach. The study outcome shows that presented system offered better rate gains in contrast to conventional HEVC. The problems associated with the complexity of the HEVC are also carried out by Oliveira and Alencar [56]. The authors have presented a solution where the splitting operation is carried out on coding unit in order to make the process quite faster. A classification approach using adaptive charecteristics has been used for this purpose. It can be seen that majority of the work associated with the video compression is oriented about HEVC enhancement as well as coding efficiency. From performance viewpoint, there are also many studied which has already addressed various presence of artifacts. However, there is a different form of factor that also controls the quality of the video which is called as saliency. It is essential to identify saliency from the video using computer vision and is used more widely for various video-based application e.g. identification of specific object, etc. Study in this direction has been carried out by Xu et al. [57] where a learning model is constructed for saliency detection. According to the plan of this implementation, a raw dataset of video is considered for object tracking which is subjected to different detections method of saliency features obtained from HEVC. The complete methodology is carried out on the basis of the data-driven methods. An extensive analysis has been carried out to ascertain the performance; however, the work lacks any form of extensive performance analysis. Study in the direction of performance upgradation is highly required for user experience. The work of Van et al. [58] has performed a downsizing of video file that is already encoded using HEVC. This process is assessed using different forms of machine learning process in order to understand the possible relationship between different coding data. This operation assists in predicting the behaviour of splitting the coding units in HEVC. The study outcome showed that random forest achieves better performance in contrast to other machine learning approach in terms of compression performance as well as increment in transmission rates. Apart from this implementation scheme, it is necessary to even extract the information associated with the depth of the coding unit. Although, there are various attempts in this regard, but the work carried out by Zhang et al. [59] have developed a model where constraint-based method for depth decision is carried out. According to this model, the decision of depth is transmitted between the rate distortion performance and coding complexity. This is further followed up by constructing a classifier using machine learning for reducing the complexity of depth corresponding to all the coding units. The study outcome is also found to offer maximum of 70% reduction in computational complexity.

Hence, there are various studies towards adoption of machine learning in video compression. Table 2 summarizes the existing research contribution.

Table.2 Summary of existing approaches of Video compression

Authors	Problems	Methods	Advantage	Limitation
Li et al. [37]	Compact resolution	Convolution neural network	Better signal quality	Not tested over video
Ryu and Kang [38]	Reduction of size of candidate mode in intra-prediction	Random forest, rate-distortion optimization	Minimal encoding time	Complexity associated is not estimated
Soh et al. [39]	Enhancing convolution neural network, complex artifacts	Temporal convolution neural network	Better signal quality	Highly iterative
Chen et al. [40]	Video compression	Convolution neural network	Better control on distortion	Invokes computational complexity
Jiang et al. [41]	Gradient problem in quantization	Convolution neural network	Better signal quality	Complexity associated is not estimated
Cheng et al. [43]	Video transmission	Reinforcement learning	Effective learning process	No extensive analysis

Lin et al. [44]	Transcoding from H.264 to HEVC	Naïve Bayesian	Minimal complexity of coding	Effect of video size on transcoding is not discussed
Zhu et al. [45]	Visual quality	Convolution neural network	Better signal quality	Induce network overhead
Iranfar et al. [46]	Selection of elite encoding process	Machine learning	Better signal quality	Effective of massive video streams not assessed
Gao e.t. al. [47]	Learning performance	Rate distortion, support vector regression	Increase prediction accuracy	Highly complex and iterative method
Xu et al. [48]	Complexity reduction	Convolution neural network	Addresses complexity for both intra/inter modes	Overall processing time could vary on different devices
Zhu et al. [49] [50]	Complexity reduction	Fuzzy logic, support vector machine	Generous scale of computational complexity reduction	Signal quality not improved
Wang et al. [51]	Coding efficiency	Convolution neural network	Better predictive performance	Higher dependency on training database.
Gao et al. [52]	Rate controlling	Game theory, support vector machine,	Better rate distortion performance	Channel capacity overrun has higher possibility
Liu et al. [53]	Coding efficiency	Support vector machine	60% reduction in encoding	Highly dependent on trained data
Duanmu et al. [54]	Fast mode decision	Machine learning	40% complexity reduction	Highly complex and iterative method
Van et al. [55]	Video coding(adaptive)	Machine learning, correlation	Better rate gain	No extensive analysis
Van et al. [58]	Analysis of performance of machine learning	Downsizing encoded video	Random forest is seen with best performance	It applies on offline videos only
Zhang et al. [59]	Allocation of Complexity	Depth of coding unit, classifier	Good complexity reduction	No benchmarking

IV. OPEN END RESEARCH ISSUES

From the prior section, it is now known that machine learning techniques has been attempted by various authors to solve various associated problems of video compression. By reviewing the existing research studies, it can be realized that video compression performance can be improved using machine learning techniques. The prime reason behind this are that machine learning can be utilized in constructing novel computational models with efficient coding schemes. All the existing system has been carried out using similar strategy where the problem associated with the compression is presented initially followed by appropriate model design to solve the issue. However, there are also associated challenges mainly in machine learning approach:

A. Frequently used scheme

A closer look into the existing system shows that convolutional neural network occupies 99% of the machine learning scheme towards video compression. A closer look into the methodology deployed shows that adoption of convolution neural network does offer some coding efficiency as the videos are normally taken from dataset and their behaviour is highly static. However, in case of untrained or dynamic video, this form of machine learning approach is incapable of performing encoding operation of the dynamic traits (e.g. orientation/position) of an object when they carry out prediction operation. Unfortunately, all the prior information of internal data (e.g. orientation/position) is almost lost. All the current information of coding units is forwarded to same neuron that obviously is overburden by massive information. This problem is not much encountered in existing system as all the works have considered a specific finite dataset, such problems only appears when large scale of compression is required to be carried out on video. Another potential problem in existing scheme of machine learning is that it is too much dependent on database. Without quantified database, it is not possible for convolution neural network to perform classification with respect to certain components. Dynamic streams of video could possibly have some artifacts, which cannot be addressed by existing frequently used machine learning scheme. Apart from this, the bigger challenging thing is tuning of hyper parameters are highly complex task and is also dependent on large database.

There is attempt towards standardization of the net weights of existing machine learning operation. This is not a big problem for similar type of feature; however, if different features are considered for study, the present model fails. Finally, the potential problem in deep learning scheme or convolution neural network is that they cannot be applied on online video compression application as the processing time of convolution neural network is abnormally very high. Unfortunately, this fact is theoretically known and there was no study or investigation towards proving this fact. Therefore, existing scheme of machine learning will require a massive and large dimensional change in order to make it suitable for performing quality and efficient video compression operation.

Problems with Internal Training Methodology: It was seen that 99% of existing approaches of machine learning uses deep learning or convolution neural network, while rest score is for support vector machine or random forest mechanism. Adoption of such problems is associated with training process issues which was never found to be addressed in any existing system. There are various parameters for performing either improvement or optimization of video compression performance. Different studies use different techniques but majority of all the techniques follow similar process i.e. it utilizes information associated with rate distortion in order to find the final outcome of enhancement. Unfortunately, such schemes are only applicable when a static size of dataset of video or video sequences is used as there are possibilities of higher scale of complications when maximized training data is considered. Normally such adoption is done to prevent problems of over-fitting but it should be known that machine learning techniques adopted till date are never meant for large datasets. In order to carry out video compression, it is required that existing approaches to consider the complexity factor of the encoder, which is not found to be considered in any existing approaches. There are good possibilities that encoder complexity could be fairly high while applying deep learning mechanism. Apart from that involvement of cost modeling could also offer more insight to the training behaviour while compression is carried out. However, none of the existing system is even reported to adopt such strategy as if complexity of encoder increases the cost will also increase. Hence, it is computationally impossible to enhance the existing training scheme to involve cost modeling. However, the studies that uses random forest mechanism is found to append a function (in order to obtain better form of internal state of the encoder and the data that is influenced by the optimization process). The decomposition of the optimization issue is carried next when bigger problem is split into smaller problem and local solutions are obtained. All the local solutions are then combined and subjected to an iterative operation. This iterative operation in random forest as well as in support vector machine is only expected to offer an elite convergence point; however, there are also fair chances of its failure too. This problem could possible evolve due to the mathematical structure of it; however, there are various reasons that leads the encoder to fail e.g. subset of training, training dataset, over-fitting, etc. For an example, if there is a case of over-fitting for one predictor in one iteration than the encoder will allocate a maximized overhead in its selection in the consecutive iteration and will therefore render worse scenario of over-fitting problems in video compression. So, inspite of looking for converging point, the system actually looks for diverging point and good solution is never obtained.

B. Inappropriate Selection of Training Data

In any case of machine learning strategy of existing system, there was never any form of discussion about the specification of the training data in video compression. We argue that selection of training data is a complex task which has never been realized in any existing literature. The prime reason behind this are – i) generally speaking a data may not be expected to be built of varied samples of video and this leads to the concept that there are various factors and data segments that are required to be considered while performing video compression. It is necessary for the encoder to obtain maximum information of the video in steady as well as in dynamic state, ii) there may be a critical requirement of using same encoder for diversified video features that could represent different quality factor of video e.g. resolution, signal quality, size, error, etc., iii) there could be possibility of presence of video dataset which is highly subjective in its nature. We strongly believe and didn't come up with any evidence in existing system when any researchers have ever considered all the above three points before performing selection of training dataset. Hence, training carried out on the video dataset without the above three-point consideration is contextually incorrect.

C. Problems of Partitioning Data & Classification

HEVC is one of the dominant video encoding systems adopted by almost all the researcher in existing system owing to its beneficial points of compression as well as its continuation in futuristic devices. However, when HEVC is adopted for enhancement (as done by almost all the authors), there was no reported case of any attempt towards exploring the potential predictors. All the enhancement approaches on HEVC are highly adhoc in nature and there is no assurity of its performance in future. Another normal approach for this problem utilized in existing system is to initiate from the natural concept and then attempt to optimize the prior solution. Again, different set of studies uses arbitrary solution followed by different enhancement approaches and then selects the best outcome. Although, some approaches offer good outcome but there is no practical or empirical justification behind the outcome, which is mathematically error-prone. It will also mean that existing approaches are yet to be validated and no reliability score is given to it.

D. Lack of Standard Models

A closer look into the existing system will show that they are all most recently deployed work towards video compression using machine learning. Although, each paper has significant information about its outcome, but there is no reported case where the outcome of the presented system is actually benchmarked with some other frequently used system. Apart from this, there are also implementations with mathematical modeling; however, there is no evidence of any convergence testing of the mathematical modeling to prove its effectiveness. None of the existing outcomes are validated with respect to processing time and storage complexity, which is another biggest pitfall of existing system.

Therefore, all the above-mentioned problems are yet unsolved and there is a need of emergent attention towards this problem. Evolving up with a computational model of machine learning to improvise the video compression process is yet a computationally challenging task.

V. CONCLUSION & FUTURE WORK

From the discussion present in this paper, there are various facts that are found viz. i) the existing approaches for image compression cannot be directly applied on all the video files as in that case the temporal charecteristics are missing in image compression approach. Therefore, a dedicated video compression is highly essential to be evolved, which is not much present in current times. ii) not all machine learning approaches are applicable directly on videos without identifying a definitive predictor in it. Deep learning, random forest, convolution neural network, etc are good algorithms but their applicability is assessed only on dataset without considering the selection problems of trained dataset based on the subjectivity of it, iii) existing system of video compression using machine learning are void of any form of benchmarking and comparative analysis with respect to processing capability, processing time, etc.

However, at the same time, it can be also realized that HEVC is the upcoming standard, irrespective of its problems. Therefore, our future direction of the work will be to achieve the following objectives viz. i) to develop a computational model where a video compression will be carried out by enhancing existing HEVC standard, ii) to extend the same computational model towards evolving up with a novel machine learning model for ensuring better compression performance.

REFERENCES

1. Al-Begain, Khalid, Ali, Ashraf, Multimedia Services and Applications in Mission Critical Communication Systems, IGI Global, 2017
2. Alexis Joly, Stefanos Vrochidis, Kostas Karatzas, Ari Karppinen, Pierre Bonnet, Multimedia Tools and Applications for Environmental & Biodiversity Informatics, Springer, 2018
3. Andy Beach, Aaron Owen, Video Compression Handbook, Peachpit Press, 2018
4. K. N. Satone, A. S. Deshmukh and P. B. Ulhe, "A review of image compression techniques," 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, 2017, pp. 97-101.
5. A. Singh and K. G. Kirar, "Review of image compression techniques," 2017 International Conference on Recent Innovations in Signal processing and Embedded Systems (RISE), Bhopal, 2017, pp. 172-174.
6. Muin, Fathiah & Gunawan, Teddy & Kartiwi, Mira & Elsheikh, Elsheikh. (2017). A review of lossless audio compression standards and algorithms. AIP Conference Proceedings. 1883. 020006. 10.1063/1.5002024.
7. J. Lee and T. Ebrahimi, "Perceptual Video Compression: A Survey," in IEEE Journal of Selected Topics in Signal Processing, vol. 6, no. 6, pp. 684-697, Oct. 2012.
8. D. J. McDuff, E. B. Blackford and J. R. Estepp, "The Impact of Video Compression on Remote Cardiac Pulse Measurement Using Imaging Photoplethysmography," 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), Washington, DC, 2017, pp. 63-70.
9. A. Pande, P. Mohapatra and J. Zambreno, "Securing Multimedia Content Using Joint Compression and Encryption," in IEEE MultiMedia, vol. 20, no. 4, pp. 50-61, Oct.-Dec. 2013.
10. M. M. Shaiboun and M. Shaheen, "Streaming medical images using video compression," 2016 IEEE Symposium on Computers and Communication (ISCC), Messina, 2016, pp. 125-128.
11. H.R. Wu, K.R. Rao, Digital Video Image Quality and Perceptual Coding, CRC Press, 2017
12. Kevin Jolly, Machine Learning with scikit-learn Quick Start Guide: Classification, regression, and clustering techniques in Python, Packt Publishing Ltd, 2018
13. S. Masuda, K. Ono, T. Yasue and N. Hosokawa, "A Survey of Software Quality for Machine Learning Applications," 2018 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW), Vasteras, 2018, pp. 279-284.
14. A. Vinothini ; S. Baghavathi Priya, "Survey of machine learning methods for big data applications", IEEE-International Conference on Computational Intelligence in Data Science(ICCIDS), 2017
15. Yang Xin ; Lingshuang Kong ; Zhi Liu ; Yuling Chen ; Yanmiao Li ; Hongliang Zhu ; Mingcheng Gao, "Machine Learning and Deep Learning Methods for Cybersecurity", IEEE Open Access, vol.6, 2018
16. Stamp, Mark. (2018). A Survey of Machine Learning Algorithms and Their Application in Information Security: An Artificial Intelligence Approach. 10.1007/978-3-319-92624-7_2.
17. Raouf BoutabaEmail authorMohammad A. SalahuddinNoura LimamSara AyoubiNashid ShahriarFelipe Estrada-Solano, "A comprehensive survey on machine learning for networking: evolution, applications and research opportunities", Springer-Journal of Internet Services and Applications, 218
18. Ilseyar Alimova ; Elena Tutubalina ; Julia Alferova ; Guzel Gafiyatullina , "A Machine Learning Approach to Classification of Drug Reviews in Russian", Ivannikov ISPRAS Open Conference (ISPRAS), 2017
19. Abhilasha Singh Rathora, AmitAgarwal, PreetiDimri, "Comparative Study of Machine Learning Approaches for Amazon Reviews", Procedia Computer Science, Volume 132, 2018, Pages 1552-1561
20. Andy Beach, Real World Video Compression, Pearson Education, 2010
21. Richard Zurawski, The Industrial Information Technology Handbook, CRC Press, 2018
22. Muhammad Vandestra, Dragon Promedia Studio, Free Opensource Video Editor For Beginner & Youtube Creator, Dragon Promedia, 2018
23. Guilherme Corrêa, Pedro Assunção, Luciano Agostini, Luis A. da Silva Cruz, Complexity-Aware High Efficiency Video Coding, Springer, 2015
24. K.R. Rao, J.J. Hwang, D. N. Kim, High Efficiency Video Coding and Other Emerging Standards, River Publishers, 2017
25. Vivienne Sze, Madhukar Budagavi, Gary J. Sullivan, High Efficiency Video Coding (HEVC): Algorithms and Architectures, Springer, 2014
26. Charles Poynton, Digital Video and HD: Algorithms and Interfaces, Elsevier, 2012
27. D. Sculley, "Compression and Machine Learning: A New Perspective on Feature Space Vectors", ACM-Proceedings of the Data Compression Conference, IEEE Computer Society Washington, 2006
28. Vinod Chandra S.S., Anand Hareendran S., "Artificial Intelligence And Machine Learning", PHI Learning Pvt. Ltd, 2014
29. Zsolt Nagy, Artificial Intelligence and Machine Learning Fundamentals: Develop real-world applications powered by the latest AI advances, Packt Publishing Ltd, 2018
30. Suresh Samudrala, Machine Intelligence: Demystifying Machine Learning, Neural Networks and Deep Learning, Notion Press, 2019
31. A. Mayr et al., "Potentials of machine learning in electric drives production using the example of contacting processes and selective magnet assembly," 2017 7th International Electric Drives Production Conference (EDPC), Würzburg, 2017, pp. 1-8.
32. R. Bhardwaj, A. R. Nambiar and D. Dutta, "A Study of Machine Learning in Healthcare," 2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC), Turin, 2017, pp. 236-241.
33. B. Shamsaei and C. Gao, "Comparison of some machine learning and statistical algorithms for classification and prediction of human cancer type," 2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI), Las Vegas, NV, 2016, pp. 296-299.
34. Y. Xin et al., "Machine Learning and Deep Learning Methods for Cybersecurity," in IEEE Access, vol. 6, pp. 35365-35381, 2018.
35. L. Bing and S. Yuliang, "Prediction of User's Purchase Intention Based on Machine Learning," 2016 3rd International Conference on Soft Computing & Machine Intelligence (ISCMI), Dubai, 2016, pp. 99-103.

36. X. Yan, H. Huang, Z. Hao and G. Li, "A human-machine cooperative approach for combinatorial optimization problem," 2016 International Conference on Machine Learning and Cybernetics (ICMLC), Jeju, 2016, pp. 838-843.
37. Y. Li, D. Liu, H. Li, L. Li, Z. Li and F. Wu, "Learning a Convolutional Neural Network for Image Compact-Resolution," in IEEE Transactions on Image Processing, vol. 28, no. 3, pp. 1092-1107, March 2019.
38. S. Ryu and J. Kang, "Machine Learning-Based Fast Angular Prediction Mode Decision Technique in Video Coding," in IEEE Transactions on Image Processing, vol. 27, no. 11, pp. 5525-5538, Nov. 2018.
39. J. W. Soh et al., "Reduction of Video Compression Artifacts Based on Deep Temporal Networks," in IEEE Access, vol. 6, pp. 63094-63106, 2018.
40. T. Chen, H. Liu, Q. Shen, T. Yue, X. Cao and Z. Ma, "DeepCoder: A deep neural network based video compression," 2017 IEEE Visual Communications and Image Processing (VCIP), St. Petersburg, FL, 2017, pp. 1-4.
41. F. Jiang, W. Tao, S. Liu, J. Ren, X. Guo and D. Zhao, "An End-to-End Compression Framework Based on Convolutional Neural Networks," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 28, no. 10, pp. 3007-3018, Oct. 2018.
42. M. Okade, G. Patel and P. K. Biswas, "Robust Learning-Based Camera Motion Characterization Scheme With Applications to Video Stabilization," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, no. 3, pp. 453-466, March 2016.
43. B. Cheng, J. Yang, S. Wang and J. Chen, "Adaptive Video Transmission Control System Based on Reinforcement Learning Approach Over Heterogeneous Networks," in IEEE Transactions on Automation Science and Engineering, vol. 12, no. 3, pp. 1104-1113, July 2015.
44. H. Lin, X. He, L. Qing, S. Su and S. Xiong, "Machine learning-based H.264/AVC to HEVC transcoding via motion information reuse and coding mode similarity analysis," in IET Image Processing, vol. 13, no. 1, pp. 34-43, 10 1 2019.
45. L. Zhu, Y. Zhang, S. Wang, H. Yuan, S. Kwong and H. H. -. Ip, "Convolutional Neural Network-Based Synthesized View Quality Enhancement for 3D Video Coding," in IEEE Transactions on Image Processing, vol. 27, no. 11, pp. 5365-5377, Nov. 2018.
46. A. Iranfar, M. Zapater and D. Atienza, "Machine Learning-Based Quality-Aware Power and Thermal Management of Multistream HEVC Encoding on Multicore Servers," in IEEE Transactions on Parallel and Distributed Systems, vol. 29, no. 10, pp. 2268-2281, 1 Oct. 2018.
47. W. Gao, S. Kwong, Q. Jiang, C. Fong, P. H. W. Wong and W. Y. F. Yuen, "Data-Driven Rate Control for Rate-Distortion Optimization in HEVC Based on Simplified Effective Initial QP Learning," in IEEE Transactions on Broadcasting, vol. 65, no. 1, pp. 94-108, March 2019.
48. M. Xu, T. Li, Z. Wang, X. Deng, R. Yang and Z. Guan, "Reducing Complexity of HEVC: A Deep Learning Approach," in IEEE Transactions on Image Processing, vol. 27, no. 10, pp. 5044-5059, Oct. 2018.
49. L. Zhu, Y. Zhang, S. Kwong, X. Wang and T. Zhao, "Fuzzy SVM-Based Coding Unit Decision in HEVC," in IEEE Transactions on Broadcasting, vol. 64, no. 3, pp. 681-694, Sept. 2018.
50. L. Zhu, Y. Zhang, Z. Pan, R. Wang, S. Kwong and Z. Peng, "Binary and Multi-Class Learning Based Low Complexity Optimization for HEVC Encoding," in IEEE Transactions on Broadcasting, vol. 63, no. 3, pp. 547-561, Sept. 2017
51. T. Wang, M. Chen and H. Chao, "A Novel Deep Learning-Based Method of Improving Coding Efficiency from the Decoder-End for HEVC," 2017 Data Compression Conference (DCC), Snowbird, UT, 2017, pp. 410-419
52. W. Gao, S. Kwong and Y. Jia, "Joint Machine Learning and Game Theory for Rate Control in High Efficiency Video Coding," in IEEE Transactions on Image Processing, vol. 26, no. 12, pp. 6074-6089, Dec. 2017
53. X. Liu, Y. Li, D. Liu, P. Wang and L. T. Yang, "An Adaptive CU Size Decision Algorithm for HEVC Intra Prediction Based on Complexity Classification Using Machine Learning," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, no. 1, pp. 144-155, Jan. 2019.
54. F. Duanmu, Z. Ma and Y. Wang, "Fast Mode and Partition Decision Using Machine Learning for Intra-Frame Coding in HEVC Screen Content Coding Extension," in IEEE Journal on Emerging and Selected Topics in Circuits and Systems, vol. 6, no. 4, pp. 517-531, Dec. 2016.
55. X. HoangVan, J. Ascenso and F. Pereira, "Adaptive Scalable Video Coding: An HEVC-Based Framework Combining the Predictive and Distributed Paradigms," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 27, no. 8, pp. 1761-1776, Aug. 2017.
56. J. F. de Oliveira and M. S. Alencar, "Online learning early skip decision method for the HEVC Inter process using the SVM-based Pegasos algorithm," in Electronics Letters, vol. 52, no. 14, pp. 1227-1229, 7 7 2016.
57. M. Xu, L. Jiang, X. Sun, Z. Ye and Z. Wang, "Learning to Detect Video Saliency With HEVC Features," in IEEE Transactions on Image Processing, vol. 26, no. 1, pp. 369-385, Jan. 2017.
58. L. P. Van, J. De Praeter, G. Van Wallendael, J. De Cock and R. Van de Walle, "Performance analysis of machine learning for arbitrary downsizing of pre-encoded HEVC video," in IEEE Transactions on Consumer Electronics, vol. 61, no. 4, pp. 507-515, November 2015.
59. Y. Zhang, S. Kwong, X. Wang, H. Yuan, Z. Pan and L. Xu, "Machine Learning-Based Coding Unit Depth Decisions for Flexible Complexity Allocation in High Efficiency Video Coding," in IEEE Transactions on Image Processing, vol. 24, no. 7, pp. 2225-2238, July 2015.

AUTHORS FROFILE

Veena S.K, She has completed her B.E in Electronics and Communication Engineering from VTU, M.Tech in CAID from VTU, India. She is pursuing her PhD from VTU, India. She has total professional experience is 09 years. She has published around 2 papers in international journal and conferences.

Mahesh K Rao, has completed his B.E in E&C from University of Mysore, SJCE, Karnataka, India. He has done M.A.Sc from University of Windsor, Ontario, Canada. He has completed his PhD from University of Wyorning, Laramine, USA. His research interest was Wireless Sensor Networks, and its applications, Image processing. applications in various fields such as for defect identifications. He has around 28 years of experience in professional.