A Robust Algorithm of Digital Video Watermarking using ELM

Shital Gupta, Megha Kamble

LNCT University, Bhopal, India

Corresponding author: Megha Kamble, Email: meghak@lnct.com

Image watermarking for copyright protection has become a widely studied subject with the sprawl of pirated content. A color, a monochrome – gray scale or a binary may be a watermark. The insertion of watermarks may be performed in a video domain unencoded or encoded. A robust technique of digital video watermarking using machine learning approach is based on Extreme Learning Machine (ELM) is proposed. Using the properties of hybrid transformations, robust features and robust zero watermarks can be extracted from videos can be built. Experimental findings show that the proposed algorithm is robust in high-efficiency video coding attacks with various parameters of quantization due the fast processing of frames. Regular image processing assaults, geometric attacks, frame-based attacks, and hybrid attacks can all be thwarted by this approach. Comparatively, the suggested video watermarking method can more accurately and completely recreate watermarking pictures.

Keywords: ELM (Extreme Learning Machine), Video Watermarking, hybrid transformations, Geometric attack.

2021. In Rahul Srivastava & Aditya Kumar Singh Pundir (eds.), *New Frontiers in Communication and Intelligent Systems*, 649–659. Computing & Intelligent Systems, SCRS, India. https://doi.org/10.52458/978-81-95502-00-4-66

1 Introduction

The Video is basically the arrangement of many frames, and every frame is shown as a picture still. Many image watermarking procedures can also be used for video watermarking. Multimedia document watermarking in video formats can take many methods of watermarking such as image and audio watermarking. The robust compression, geometric changes/cutting frames, precision of frames coding without visual artifacts are all the things to consider in video watermarking and should be attentive to the operating time or speed of the produced video. The video watermark is embedded in a moving or moving part. The watermarking is composed of two forms, the same watermark and the different watermark. The embedding of the watermark in the frameless watermarking portion uses the same watermark in order to prevent removal of watermarks by an unauthorized user. Frame moving watermarking is a technique for inserting information into an audio file to prevent the other people from making any additional information visible. Watermark is a good way to ensure that data that only the owner knows [7] protects against media copyright. The compression performance of H.264/advanced video coding (AVC) can no longer meet the coding requirement due to the popularity of full-high-definition (FHD) and ultrahigh-definition (UHD) videos. Since H.265/high efficiency video coding is the current standard, it has gradually superseded AVC as a widespread standard for video compression. Meanwhile, widespread HEVC video piracy is wreaking havoc on the entertainment business and must be addressed immediately. With robust video watermarking as one of the information hiding methods, the watermark is included in the video to provide copyright protection. Imperceptibility and robustness are the two most important measures for evaluating the performance of a robust video watermarking system. To say a watermark is imperceptible is to say that it cannot be seen by the human eye and is difficult to detect by detecting equipment. A video's quality won't be adversely affected by the watermark's insertion. Robustness refers to a watermark's capacity to withstand attacks, such as typical signal processing attacks or malicious operations, without removing the embedded watermark. Video watermarking necessitates real-time performance due to the fact that the watermark's embedding and extraction speeds must not be slower than the video's frame rate. The bit rate increase is measured by the bit increase ratio (BIR). Watermarking is also evaluated on the basis of its capacity. Watermarking videos is less concerned with capacity than it is with the sheer volume of frames in the video [5].

2 Related work

In Paper [1], an invisible digital watermark technique based on DWT and DCT domains is proposed.According to the algorithm's test results, watermarked videos had PSNR values as high as 37 dB when watermarking was done correctly. The proposed algorithm has been shown to be robust against compression of the HEVC stream. Quantification based embedding was also suggested for video watermarking on the DCT domain but in three dimensions (3DDCT). A second analysis in [2], used uniformly the algorithm for image insertion in each bit plane, quantification decomposition into 8 bits coefficients of selected DCT Blocks, taken from the video original. An additional study proposed a pattern recognition algorithm. Digital data embedded in the DCT response AC coefficients. Singular value decomposition is used in this paper[3] to provide a new fast and resilient video watermarking approach for RGB uncompressed AVI video sequences in the DWT domain (SVD). Scene change detection is carried out for the purpose of embedding. The LL3 sub-band coefficients of the video frames contain the singular values of a binary watermark. A high-quality signed video has been produced as a result. Six distinct video processing procedures are tested to see if the suggested method is robust enough. The high computed PSNR values suggest that the video's visual quality is excellent. There is a strong correlation between the extracted watermark and the implanted watermark based on

the low bit error rate and the high normalised cross correlation values. For real-time use, the suggested scheme's temporal complexity has been evaluated. The proposed algorithm's embedding and extraction have been shown to be well-optimized, according to this study. The technique is dependable and outperforms other recently published methods of a similar nature.On a piece of paper During network transmission, the watermarked video will be subjected to multiple attacks, including transcoding, noise, and temporal synchronisation. No consideration is given to combined attack analysis or testing in the current algorithmic frameworks. This research provides a new approach based on a finite state machine and the DTCWT-SVD transformation for correctly extracting the watermark from video during network propagation. ZigZag first scans the watermark image into numerous segments, and then the carrier video is split into multiple groups of frames. DTCWT low frequency coefficient matrix singular value changes are used to implant each grouping of watermarks into the comparing video outline, a finite state machine generates this. According to a predetermined relationship between individual values, we then extract the watermark. For example, noise, temporal synchronisation, and other attacks on the proposed approach were successfully thwarted in the experiments. The watermark information in the video may also be precisely extracted after network propagation, so that the video can successfully resist the combined attacks that it faces during network propagation. As a result of this research, This study proposes a new watermarking method for copyright protection of multimedia colour films [5]. The innovation in the described method is the invention of a hybrid DWT and DCT-based digital video watermarking of colour watermark logos employing index mapping technology. The PSNR and SSIM are used to evaluate the watermark's distortion and its ability to withstand various forms of attacks, whereas Stir Mark is used to evaluate its resilience. The suggested video watermarking technique offers enhanced imperceptibility and robustness against signal processing attacks, both of which are compatible with the visual system of the human eye.When it comes to pen and paper Digital watermarking is a copyright-protection and authentication method that utilises data obfuscation. Digital watermarking on DWT and DCT domains is provided with an invisible watermarking algorithm. A video stream was encoded with a binary watermark image included in the middle sub band coefficients. Watermarked videos have PSNR values as high as almost 37 dB with the optimal watermarking strength, according to experimental results. The suggested approach was shown to be resistant to the HEVC stream compression standard, which was used to test its strength. Hence, its potential for use in copyright protection and authentication is high.Discrete Wavelet Transform is used in paper [7] to insert a grey image from Arnold Transform into a movie (DWT). A histogram-based scene change technique is used to authenticate the video with various regions of the watermark in the proposed scheme. Each frame is broken down into three separate planes, making it easier to see. In order to decompose into smaller bands, DWT is applied to a single plane in each frame. Each of the 8-bit planes of the concealed image is divided into two. An Arnold transform further muddles the bit plane image for enhanced watermark security. Video quality is unaffected because Embedding is carried out in the DWT high frequency coefficients as well as mid frequency coefficients. The inverse processing stages are used to extract the hidden picture from the marked video. Testing for robustness involves putting a marked video through a slew of video and image processing assaults. Frame averaging and frame dropping are not a problem because of the suggested system, according to the findings of the simulations.Video watermarking is described in detail in article [8] with the goal of overcoming issues with security. There have been a number of ways recently proposed to meet security and signature-invisibility are only a few of the additional constraints placed on video apps.It is proposed in this study that feature regions be used as the basis for a novel video watermarking approach. This strategy is unique in that it makes use of community sourcing to identify feature regions. Before anything further happens, a video summary is created. Using this summary, the initial feature regions can be identified through crowdsourcing. When it comes to detecting the second sort of feature region that is browsed by moving objects, a mosaic is generated from original footage. In the final step, the signature is incorporated into the mosaic formed by combining these two types of feature

regions using the multi-frequential watermarking strategy. As a result of the wise selection of the embedded target, experimental results reveal a high level of invisibility. This method is also resistant to a wide range of attacks; including collusion attacks. In the paper [9] Data on the internet is well-protected with the use of watermarking, steganography, and cryptography, all of which are widely employed. For copyright protection, tamper detection, authentication, locating unauthorised copies, and as information-carrying watermarks, watermarking is utilised. Multimedia applications include images, music, video, and all forms of text. It has been discovered that LSB (Least Significant Bit) watermarking can be utilised to protect spatially vulnerable data. It's simple to put into action, and it works well. DCT and DST-LSB (discrete sine and cosine transform) audio watermarking systems with substantially higher sensitivity than LSB have been developed. Tamper detection has been detected using PSNR (Peak signal to noise ratio), as well as NC (Normalized correlation). The proposed approaches are only capable of detecting tampering, but not of locating it.When it comes to digital computer technology, there is a growing demand for video watermarking products because of the abundance of free videos available online that need to be safeguarded and verified for copyright and ownership purposes. The need to safeguard digital media and intellectual property rights has grown significantly. As a result, digital watermarking is the most effective method for protecting audiovisual items from unlawful transactions. There has been an increase in the adoption of digital watermark technology as a result of its convenience. Information about the copyright holder of this digital item is contained in a watermark. Any digital information that describes the copyright, such as a brand image, a serial number, or any other digital information. As an added measure of security, the watermark and digital interface can be combined in this instance. Whenever it's necessary to show copyright ownership, the watermark is removed from the digital interface. It is necessary that the watermark be resistant to a variety of attacks, both deliberate and unintentional[11]. In order for a watermark to be effective, it must have properties such as being undetectable and reversible.

3 Problem Definition

Digital video watermarking is a technique for installing hidden data or additional data into the cover picture which can later be removed or recognized for various purposes, such as permission, owner proof, content insurance, copyright security etc. Several digital picture watermarking procedures exist, and the following are their disadvantages.

The above literature survey revealed the following challenges.

(i) Multidimensional data like pictures with current digital watermarking methods are computationally complex.

(ii) It is difficult to maintain robustness and watermarking capability in video watermarking

(iii) Current video digital image watermarking techniques allow for low fidelity authentication of images

4 Proposed Work

A. Frame Extraction from the Video

Watermark embedding in video has been presented as a solution.displays a block diagram in figure 1 First, the host video is divided into incompatible M x N-size frames. Secondly, pre- sampling is performed using the frame rate to minimize the number of frames to be analysed. The main frames are next extracted using the colour histogram difference calculation from the collection of pre-sampled

frames. Afterwards, the main frames are watermarked by an intense learning computer with a binary watermark.

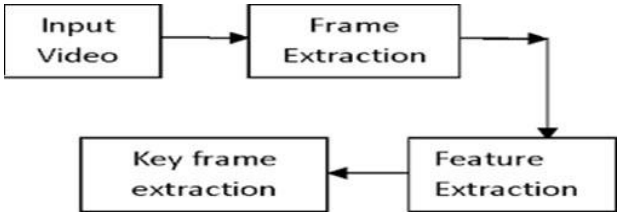

Fig. 1. Frame Extraction

B. Watermark Embedding

The watermark integration method uses a watermark key and the watermarking algorithm to generate the digital watermark image. The integration method varies depending on which image domain, e.g., space, frequency domain or wavelets is processed and shown in figure 2. In the proposed watermarking procedure, the watermark image is transformed into watermark bits and, with the watermark bits, the golden sequence is added and its combination is used to integrate the watermark in the selected block location. The watermarking procedure is as follows

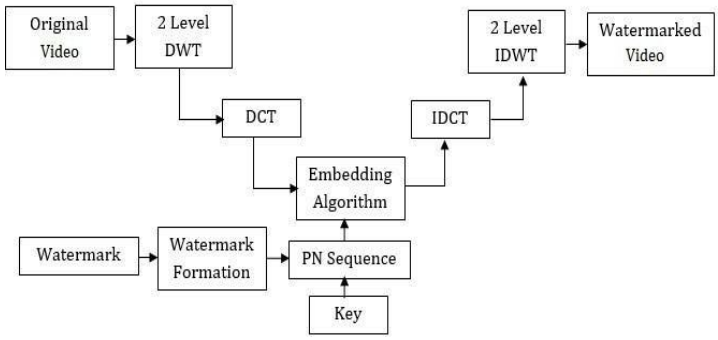

Fig. 2. Block diagram of image transmitter

Algorithm 1: Watermark embedding operation

Step 1: Extract frames from cover video and decompose each frame using DWT, LL, LH, HL, HH non overlapping sub bands.

Step 2: Apply second level DWT to LL sub band, and get four more sub bands LL1, LH1, HL1, HH1.

Step 3: Apply DCT to above blocks.

Step 4: Take watermark and convert 2D watermark to 1D form for embedding.

Step5: Pseudo randomly generate two independent sequences. There are two sequences that are utilised to embed the watermark bits 0, and 1, respectively. Pseudorandom sequences mid-band

element counts must match those of the DCT-transformed DWT sub-bands, else the pseudorandom sequences fail.

Step 6: Embed the two pseudorandom sequences, with a gain factor, in the DCT transformed 4x4 blocks of the selected DWT sub-bands of the host video.

Step 7: Apply inverse DCT Step 7: Apply IDWT

Step 8: Finally, we will get watermarked Video

C. Attack Construction

Video sequences are a collection of still images called video frames that are consecutive and are spaced in time. Therefore, any watermarking device can be used to reinforce any frame in the video. To introduce a special video watermarking system, the watermarking designer must understand certain aspects of this media. First, video is larger than image size, so that the storage size of video is usually often codified. This means that the watermarking system must be stable video compression by design (the main compression standards are: MPEG-1, MPEG- 2, MPEG-4 part 2 and part 10). Second, the large volume of information that the user has to provide at a given time increases limitations in order to take real-time applications such as video streaming into account. The most significant thing, finally, is that temporal attacks are not found in photographs such as frame drop, transposition of the frame, insertion of the frame, Others.

D. Watermark Extraction from the Video

Extraction is simply the reverse function of the embedding and includes three steps: video pre-processing and detection with watermark extraction and post-processing with video watermarking. Next, convert the watermarked video to pictures. With the scene change check of the frame, the presence of watermark is detected. When a scene-changed frame is present, it shows "Watermark" The reverse method of embedding is an extraction. The watermark image is extracted by the subtraction process between the specific sub band of the watermarked video frame and the cover video frame. The same sub band for the embedding is selected. Property or copyright protection is demonstrated by the extraction of the watermark from, among other things, a given scene-changed watermarks.

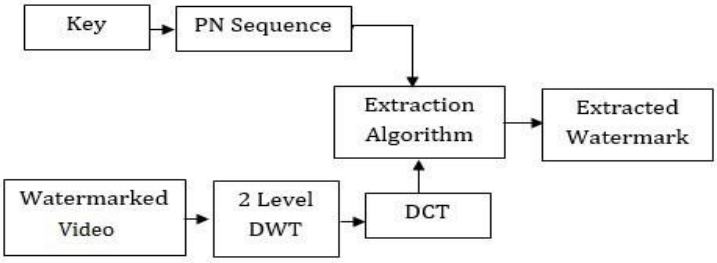

Fig. 3. Block diagram of image receiver

The quality parameters for above:

PSNR (Peak signal to noise ratio), MSE (Mean Square Error), SC (Structural Content), AD (Average Difference) NBE (Number of Bits Embedded), EC (Embedding Capacity).

Previous Work Related with This: In some previous literature, researcher uses plane DCT or simple DWT based watermarking technique. This paper proposed combined or hybrid watermarking with DCT

and DWT.

Benefits of This work:

(i) In this work we will improve the quality of the watermarked video.

(ii) PSNR increases and MSE decreases.

(iii) After implementing this work watermarked video is withstand with all kind of attacks like compression attack and noise attack.

E. Transformation

Discrete Wavelet Transform

Hierarchical picture decomposition is made possible by the Discrete Wavelet Transform (DWT) [10]. For non-stationary signal processing, it is a useful tool. Small waves, known as wavelets, of changing frequency and short duration are used in the transform. The image's frequency and spatial description are both provided by the wavelet transform. As opposed to a traditional Fourier transform, this technique retains temporal information. In order to construct a wavelet, a fixed function known as the mother wavelet must be translated and dilations. Multi-resolution image decoding is possible with DWT [11], which describes a picture at many resolutions simultaneously. High and low frequencies are separated by the DWT. Information regarding the edge components can be found in both the high-frequency and the low-frequency parts. Using high frequency components for watermarking is more common since the human eye is less sensitive to changes in the edges [13].

Discrete Cosine Transform (DCT)

By using the DCT, you may construct a finite sequence of cosine function shapes at varied oscillation frequencies. It's a technique for transforming any signal into its fundamental frequency components. It is a block-based approach that separates the image into N x N chunks that do not overlap. Because of this, the image is divided up into three distinct frequency ranges using the DCT transformation. HVS is particularly responsive to low-frequency components (Human Visual System). High frequency components are also vulnerable to a variety of attacks. A trade-off between robustness and imperceptibility can be achieved by using middle frequency components. In terms of continuity, DCT outperforms DFT. DFT is a periodic representation of the signal with truncated coefficients, whereas DCT is beneficial for representing the signal with the least number of coefficients possible. Video Watermarking Algorithm Based on DWT and DCT. DCT is a widely used standard for compressing images and video. It has the ability to compress a signal into a smaller form. Another problem with DFT is that it truncated the signal's coefficient, resulting in a loss of data. The original signal is closely approximated by DCT, which has a continuous periodic nature.

Extreme Learning Machine

Single hidden layer feed forward neural network (SLFN) architecture is used to build the Extreme Learning Machine (ELM) [16-17]. This is a departure from the typical gradient-based neural network learning algorithms. In addition to achieving the lowest possible training error, ELM also achieves the lowest possible output weight norm. ELM training is rapid, accurate, generalizable, and provides a solution in the form of a linear equation system. Because ELM has no control parameters like halting criteria, learning rate, learning epochs, etc. for particular network architecture, this network's implementation is incredibly easy. Based on a continuous probability distribution function, the weights and biases of the input and hidden layers are randomly selected. In our simulation, we use a probability distribution that is equal to one percent. In this scenario, the sigmoid function is used, which is one of many functions that can be used for learning. ELM's flow diagram is shown here (Extreme learning

machine is given below).

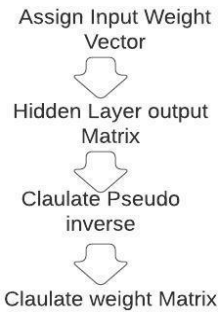

Fig. 4. Flow diagram of ELM

ELM's advantages in video watermarking include:

(i) First, it has a reasonable ability to generalize

(ii) Because the weights are chosen at random, the process is lightning-fast.

(iii) Because it must process a huge number of frames in order to develop real-time video watermarking applications, it is an excellent choice. Selected frames can be processed in milliseconds utilizing ELM technology.

5 Simulation and Result

This section summaries and discusses the main finding of work, where the evaluation of work is simulated on MATLB 20. The result achieved surpass the earlier work in this area in terms on watermarking parameters. To evaluate the performance of proposed work standard video is used in MATLAB 20 simulation software shown in figure 5 and figure 7. Figure 2, Figure 3 and Figure 4 depicts the flow diagram of proposed video watermarking scheme above. The video in the host is first divided into unconverted frames M x N in dimension. Secondly, the pre-sampling is carried out using the frame rate to minimize the number of frames to be analysed. Next, the main frames are extracted using the colour histogram difference measure from the collection of pre-sampled frames. In the last key frames is watermarked with extreme learning machine.

Robust Analysis

This section summarizes and discusses the main finding of work. The watermark used in the experiment is an image if 64*64 pixel. The video input avi format as shown in figure. In the simulation watermark is embedded using MATLAB 20. Multiple experiments were used to determine the solution's resilience. When evaluating the watermarked video's quality, we looked at its normalized correlation, peak signal-to-noise ratio, and mean square error.

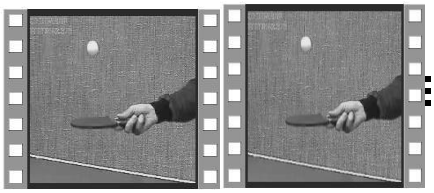

Fig. 5. Cover Video I before watermark, after watermark

Fig. 6. Watermark

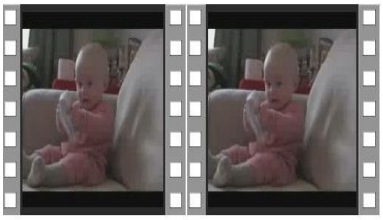

Fig. 7. Video Frame II Before and after watermark

Fig. 8. Watermark

In above figure 5 and figure 7 is cover video and figure 6 and figure 8 is watermark. Total time elapsed for Embedding the water is 7.74sec and total time spend on watermark extraction is 0.59sec. When there is attack then total time 0.54 with correlation 0.98. Table 1 with calculated parameter NC, MSE, PSNR.

Table 1. Calculated Parameter

Video	**Parameter of Used Video**		
	NC	**MSE**	**PSNR**
Video 1	**.987**	**2.199**	**61**
Video 2	**.977**	**1.432**	**58**
Video 3	**.988**	**1.179**	**63**

Table 2. State of Art Comparison

Algorithm	**SSIM**	**ΔPSNR**	**Detection Rate**	**FPR**	**Bit Rate Increase**
Proposed	**--**	**0.15**	**100%**	**0.5**	**4%**
Mareen et al	**0.966**	**0.14**	**100**	**0.4**	**3.6%**
Jiang el at	**0.989**		**78%**	**--**	**0.1%**
Chen et al	**--**	**0.06**	**76%**	**--**	**0%**
Mat et al	**--**	**1.25**	**--**	**--**	**3.6%**

The higher the PSNR score, the better the visual quality. Using the approaches proposed in table 1, the

PSNR value in several video sequences is greater than 30db. Unfortunately, the SSIM is not a perceptibility metric used in current approaches. The PSNR, or the SSIM between the watermarked video and the unwatermarked video, is often used instead. Averaging the results from each sequence, Table II displays the suggested method's PSNR and SSIM. In addition, the table compares the results with those from other, more advanced methods with a similar level of computing difficulty.

Compared with existing algorithm

Using the proposed algorithm, we may assess how well watermarks are detected in comparison to the current standard. Several attacks are compared in Table 2 to see how well the workers fared. Attacks like this one can cause up to 50% of a target's frame to be removed, known as frame dropping. Frame averaging attacks occur when the current frame is replaced by the average of two current frames. When the current frame is replaced with the frame immediately in front of it. The term "frame repeating attack" refers to an attack in which the current frame is repeatedly played. A high watermark extraction rate is seen in Table 2 even when the video drop frame is greater than 50%. This is true for both transcoding and decoding. The proper extraction of the watermark's rate is higher than that of the literature in other attacks on temporal synchronisation.

Table 3. Comparison Table

Name of Attack	**Paper**	
	guanxi et al.	**Proposed Method**
Paper salt Noise	**0.901**	**0.904**
Addition of noise	**0.971**	**0.975**
Frame swapping 25%	**0.988**	**0.990**
Frame Swapping 50	**0.988**	**0.990**
Frame Dropping 25%	**0.988**	**0.990**
Frame Dropping 50%	**0.988**	**0.990**
Frame Averaging	**0.952**	**0.954**
Frame Repeating	**0.988**	**0.990**

6 Conclusion

A robust technique for digital video watermarking technique using machine learning has been proposed in this paper. The metrics used to measure the transparency of the proposed method were based on the image quality evaluation and the proposed solution demonstrated that it also meets the transparency criteria of the watermarking method. A scenario of once encoding and multiple decoding was carried out in order to test the robustness of the proposed process. The video marked as a watermark has once been coded using the full budget for encoding. While the higher watermark scales could be rebuilt from decoded video at a higher bit rate, the experimental results showed at least a smaller scale of the watermark. The system has been found to be strong against selected attacks and is superior to another video watermarking system based on robust machine learning approach. These procedures are done easily and in milliseconds. In total, the necessary processes have been optimized and the scheme is found to be ideal for developing applications for real-time watermarking.

References

[1] Panyavaraporn, J. and Horkaew, P. K. S. T. (2018). DWT/DCT-based Invisible Digital Watermarking Scheme for Video Stream. In *10th International Conference on Knowledge and Smart Technology (KST)*, 154-157.

[2] Nguyen, T. T. and Nguyen, D. D. (2015). A robust blind video watermarking in DCT domain using even-odd quantization technique. In *ICATC*, 439-444.

[3] Goel, B. and Agarwal, C. (2013). An optimized un-compressed video watermarking scheme based on SVD and DWT. In IC3, 307-312.

[4] Guangxi, C. et al. (2019). Combined DTCWT-SVD-Based Video Watermarking Algorithm Using Finite State Machine. In *ICACI*, 179-183.

[5] Kunhu, A. et al. (2016). Index mapping based hybrid DWT-DCT watermarking technique for copyright protection of videos files. In *ICGET*, 1-6.

[6] Madhukar, B. N. and Jain, S. (2016). A Duality Theorem for the Discrete Sine Transform (DST). *International Conference on Advanced Computing and Communication Systems,* 156–160.

[7] Sujatha, C.N. and Sathyanarayana, P. (2019DWT-Based Blind Video Watermarking Using Image Scrambling Technique.). *Information and Communication Technology for Intelligent Systems. Smart Innovation, Systems and Technologies*, 106. Springer, Singapore.

[8] Kerbiche, A. et al. (2018). A robust video watermarking based on feature regions and crowd sourcing. *Multimedia Tools Application*, 77: 26769—26791.

[9] Sripradha, R. and Deepa, K. (2020). A new fragile image-in-audio watermarking scheme for tamper detection. In *ICISS*, 767-773.

[10] Kim, M., Li, D. and Hong, S. (2013). A Robust and Invisible Digital Watermarking Algorithm based on Multiple Transform Method for Image Contents. In *Proceedings of the World Congress on Engineering and Computer Science*.

[11] Swanson, M., Kobayashi, M. and Iewfik, A. (1998). Multimedia data embedding and watermarking techniques. In *Proceedings of IEEE*, 1064--1087.

[12] Wang, L. et al. (2012). Real-Time Compressed-Domain Video Watermarking Resistance to Geometric Distortions. In *IEEE Multimedia*, 19(1): 70—79.

[13] Masoumi, M. and Amiri, S. (2012). A Blind Video Watermarking Scheme Based on 3D Discrete Wavelet Transform. *International Journal of Innovation, Management and Technology*, 3(4): 487-490.

[14] Leelavathy, N., Prasad, E. V. and Kumar, S. S. (2012). A Scene Based Video Watermarking in Discrete Multiwavelet Domain. *International Journal of Multidisciplinary Sciences and Engineering*, 3(7): 12-16.

[15] Lin, M. B. et al. (2005). Fully complex Extreme Learning Machine. *Neurocomputing*, 68: 306-314.

[16] Huang, G. B. *The Matlab code for ELM is available* onhttp://www.ntu.edu.sg.

[17] Mareen, H. et al. (2018). A Novel Video Watermarking Approach Based on Implicit Distortions. In *IEEE Transactions on Consumer Electronics*, 64(3): 250-258.