

# AN INNOVATIVE APPROACH FOR DENOISING OF THE DIGITAL VIDEO STREAM

Anand B. Deshmukh<sup>1</sup> and Dr. Sanjay V. Dudul<sup>2</sup>

<sup>1</sup>Assistant Professor

Department of Information Technology

Sipna College of Engineering and Technology, Amravati (MH), India.

<sup>2</sup>Prof. and Head, Department of Applied Electronics,

Sant Gadge Baba Amravati University, Amravati (MH), India.

<sup>1</sup>abdeshmukh772@gmail.com

<sup>2</sup>svdudul@gmail.com

**Abstract:** Everyday tones of video signals are generated, transmitted and analyzed. The video contents are created for educational purpose, entertainment purpose, surveillance purpose, medical imaging purpose, weather forecasting, satellite imaging and many other significant places. During the different phases of video content preparation, transmission and analysis some unwanted signal get interfered with the true contents. Particularly, the medical imaging signals, since they are weak signals, are more prone to the unwanted interferences. Such unwanted interference of the noise signals makes it difficult to analyze the critical information in the video contents and hence, need of denoising process arises. A decent video denoising framework assures visual improvement in the video signals or it serves as the significant pre-processing step in the video processing steps like compression and analysis. Through this paper, we are about to disclose an efficient video denoising framework which takes the noisy video signal in the form of frames per second, and performs the video denoising using shot detection, compensation, intensity calculations and motion estimation process.

**Keywords:** Video denoising, PSNR, SSIM, MSSIM

## 1. INTRODUCTION

It was revealed in the press event of the YouTube that, different platform users are uploading 500 hours of video contents every minute [1], this takes us to 30,000 hours of video contents every hour and 7,20,000 hours of data per day. This means, to watch the video contents uploaded in a day, an average person needs entire lifetime. Apart from this, many video signals have never been uploaded to any social media platforms but are critically generated for education purpose for example medical imaging, weather analysis and forecasting. Some video signals are recorded and has to be reproduced repeatedly with high quality of reproduction like cinemas or personal videos recorded through mobile cameras. Some video signals are recorded from remote location and has to be transmitted to the base station like satellite videos, stream recorded through military drones or live match telecasting. At all these corners, video signals may get interfered with some unwanted signals. These unwanted signals must be removed for better reproduction of the video signal or for better analysis purpose. The methodology which is used to remove the noise signals from the video signals is called as the denoising framework.

Through the proposed title, an innovative approach for denoising of digital video stream, we are about to disclose the new denoising framework using shot detection and motion compensation technique. In the proposed denoising framework, the video stream which is to be denoised is input to the framework. Then, we can add Gaussian noise or Impulse noise to the input video frames. After this, shot is detected and motion compensation is done, subsequently, image intensity is calculated, various parameters like motion vector

descriptors, global motion attributes and new motion fields are initialized, motion estimation performed for moving region and finally, with few more critical steps the noise in the video stream is removed. The performance of the proposed denoising framework is evaluated by considering the video frames with different frame length and different frame rate. Noise is added to the mentioned video stream in step of 10%, up to 90% value. Finally, cogent analysis is carried by considering different ascendancy parameters like Peak Signal to Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM) and Mean Structural Similarity Index Measure (MSSIM). In the subsequent section of the paper, previously reported denoising frameworks are critically reviewed, research methodology of the proposed denoising framework is disclosed, discussion on the outcome with respect to the ascendancy parameters are carried out, respectively.

## 2. VIDEO DENOISING: AN EAGLE EYE

The S. Gu et. al. [2] address the issue of the information sharing between the neighboring frames. Authors handle this issue by sampling and fusing multi-level features between the adjacent frames for enhancing the video stream. To sample spots in the frame on spatial feature, learnable 2D offset is added to the 3D convolution. The effectiveness of the proposed system is verified through the interpolation, super resolution, denoising and recognition applications, where effective outcomes are observed. The subsequent figure discloses the application of the proposed system in video interpolation.

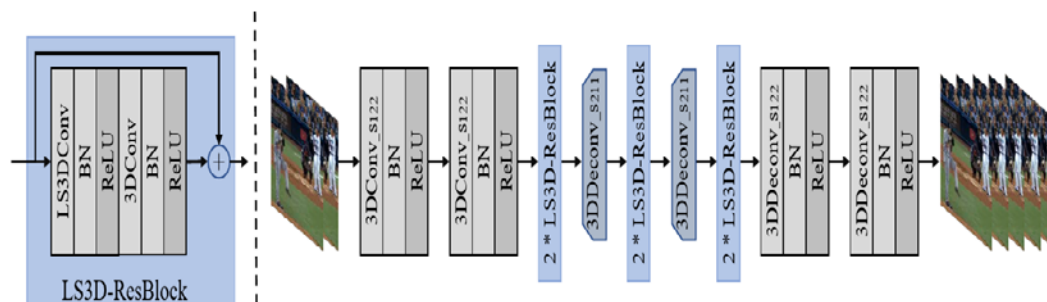


Figure 1. Proposed Blocks for Video Interpolation

Another M. Li. [3] also have been found to be working for addressing the issue of loss of critical information between the neighboring frames. Authors have proposed a video frame deblurring algorithm using denoising engine framework for restoring the blurred frames in the video stream. For implementation of this, authors are using Adaptive Laplacian Regularization in the denoising engine for restoration of the blurred video frames. For implementation of this system, a methodology is designed in which, the self-similarity redundant information is extracted using nonlocal means regularization then novel approach that is mixing of the Nonlocal Means Regularization with the denoising regularization is applied. For restoration of the videos, authors have used gradient decent approach. Different approach was disclosed for dehazing, deblurring, denoising and compression using fundamentals of the deep learning model. A. M. Tekalp et. al. in [4] have designed deep learning model for restoration and compression of the images and videos. This is assumed to be the new application of learning-based methods for such applications.

For the sake of compressing the video signals, contemporary video codecs are using correlations between the temporal and spatial planes. In case if the video frames are corrupted by any means, then it becomes difficult to compress the video signals at highest expected level. C. Chen et. al. [5] have represented use of temporal filtering technique for denoising of the video frames and further, use of said technique for lossy video compression. Through this research work outcome is like through video denoising it is possible to reduce the bit rates while maintaining the quality of the video stream.

Outcome disclosed by the author is that highest compression efficiency is observed for videos of different resolution, formats and different noise levels. The subsequent figure discloses change in PSNR value for frames with and without denoising of video stream.

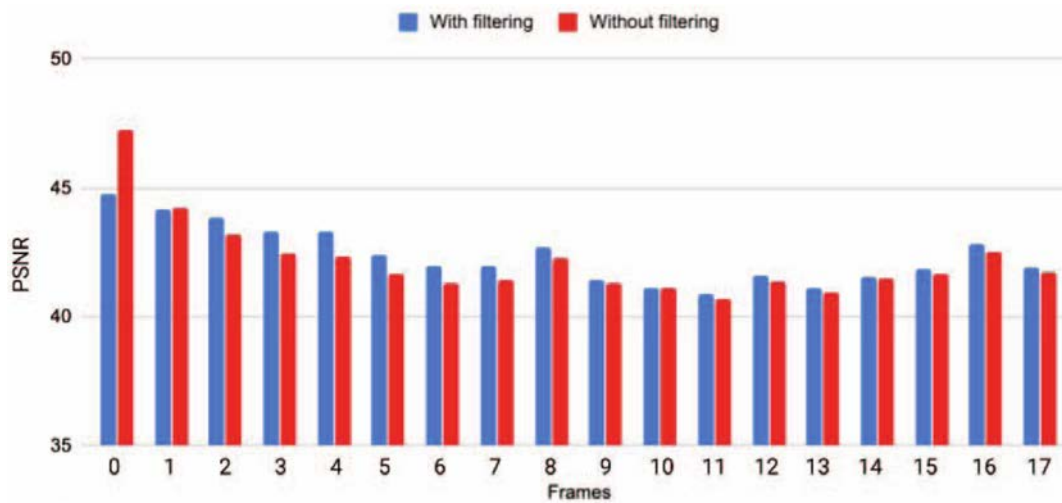


Figure 2. Change in PSNR Value with and Without Compression

J. Enhorn et. al. [6] presented a use of temporal filtering for video coding using fundamentals of bilateral filtering technique. Infact authors have designed a motion compensated temporal bilateral denoising pre-filter for video coding. Author have done the filtering process before encoding. For filtering of the frame, along with current frame, two immediate passed and two subsequent frames along with the quantization parameter are used for filtering of the frame. Using the assumed technique, authors have claimed average luma BD rate of -3.9% as compared to the VTM-7.0 encoder, in random access configuration.

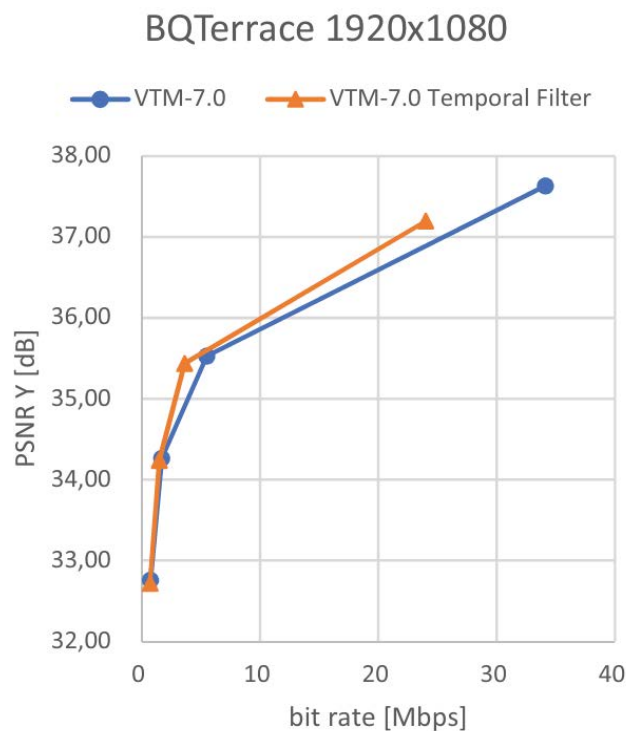


Figure 3. Rate Distortion Curve for Class B BQTerrace Sequence at Different Values of QP

Data science is coming out to be proven technique for designing of images and frames. Huanjing Yue et. al. [7] have proposed an technique for video denoising using supervised learning approach. Authors have attempted to rectify the critical issues by designing motions for controllabel objects. Subsequently, each static positions are captured significant number of times. Due to this, it is possible to create dataset with 55 groups of noisy-clean videos with ISO from 1600 to 25600. Using this dataset, denoising framework is designed using temporal, spatial and channel correlations of video frames. The Bayer patterns of the raw videos are packed into four sub-sequences that is RGBG sequences for denosing using RViDeNet. Along with the denoised video stream in raw format, sRGB format is also available at the output. The subsequent figure discloses the proposed denoising framework.

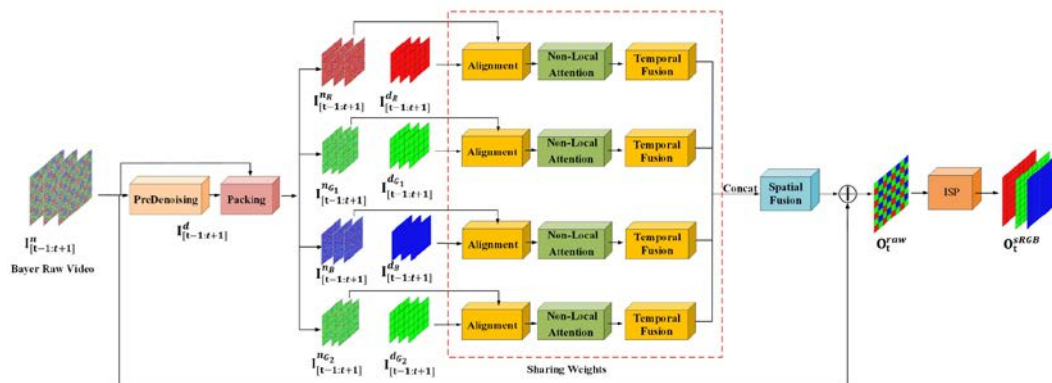


Figure 4. Proposed RViDeNet Framework

Authors find that noise to noise method needed paired noisy data also model which are trained using unpaired database gives unpredictable or limited outcomes. Further, use of the optical flow methodology for estimation is also having restricted to clean videos. To address the issues in different circumstances, Songhyun Yu et. al. [8] proposes a framework for fine tuning the video denoising and optimal flow estimation by the use of unpaired noisy frames. The subsequent figure discloses the training phase and the testing phase in the proposed methodology.

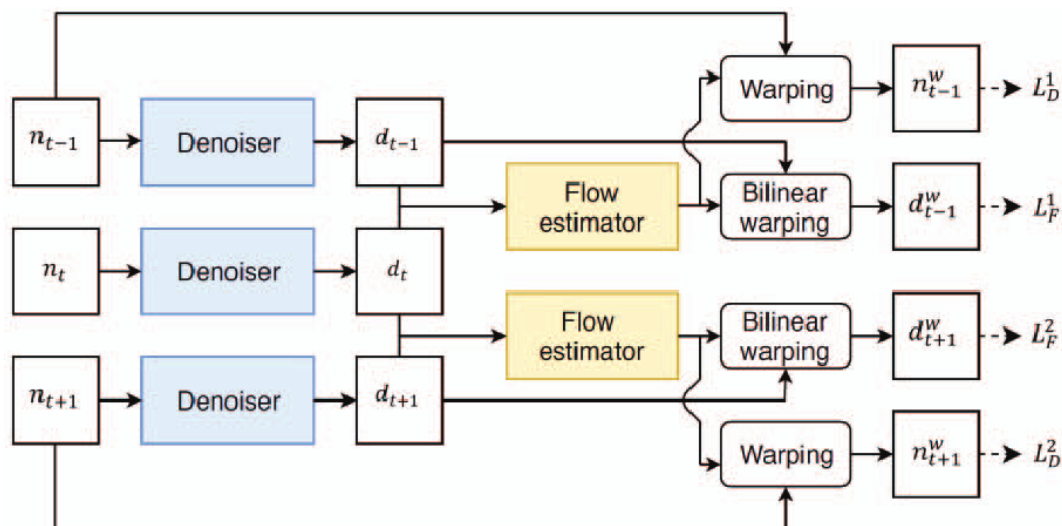
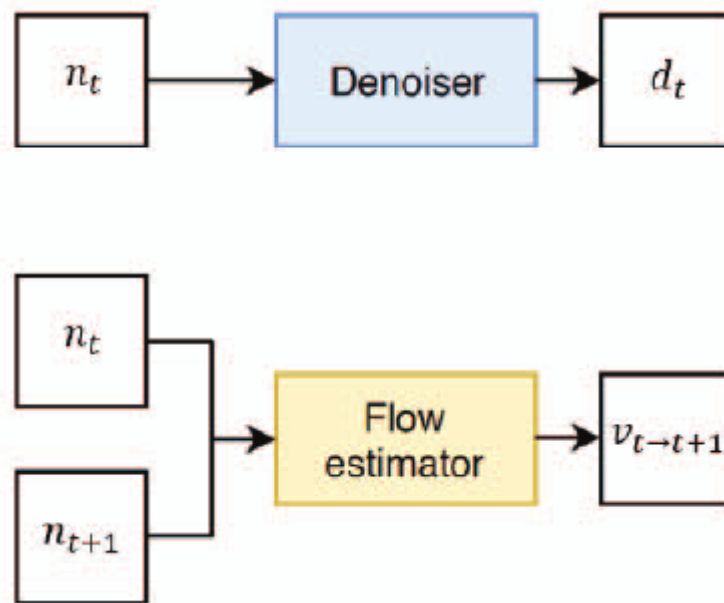


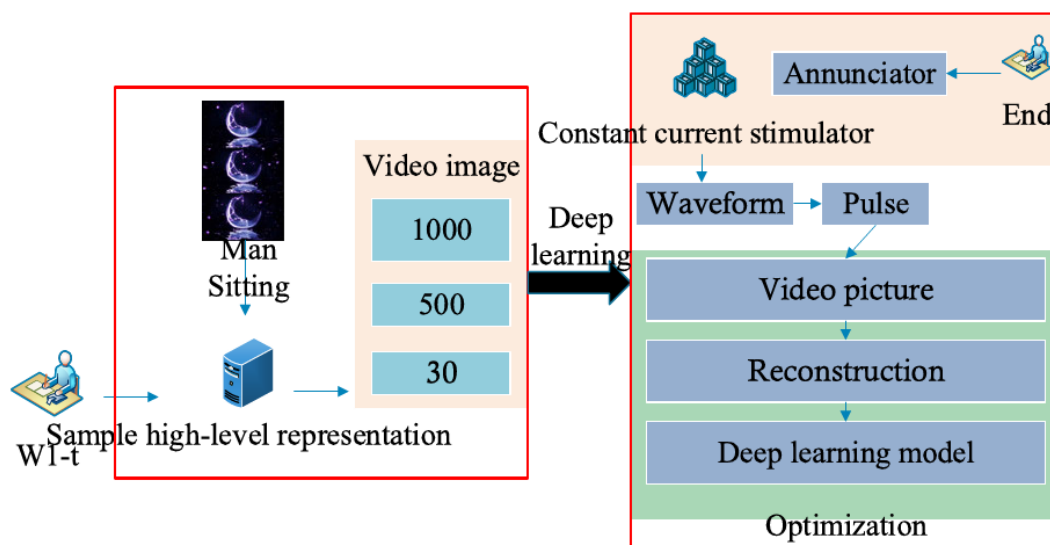
Figure 5. Training Phase of the Proposed Framework

Authors further experienced that, if the denoising performance is improved then accuracy of the flow estimation increases. Further, if improved performance is observed in flow estimation, it gives better training data set to the denoising framework. Authors further claimed that, by using the mentioned framework it is possible to get better performance in video denoising and it is also possible to get accurate optical flow in noisy video stream.



**Figure 6. Testing Phase of the Proposed Framework**

Ang Li et. al. [9] have addressed multiple issue like extracting the image features effectively, optimizing the image quality evaluation method using the deep learning approach, further combining optimal flow method and edge detection algorithm using improved optical flow estimation and deep learning technique. Subsequent figure discloses the process of deep learning model formation.

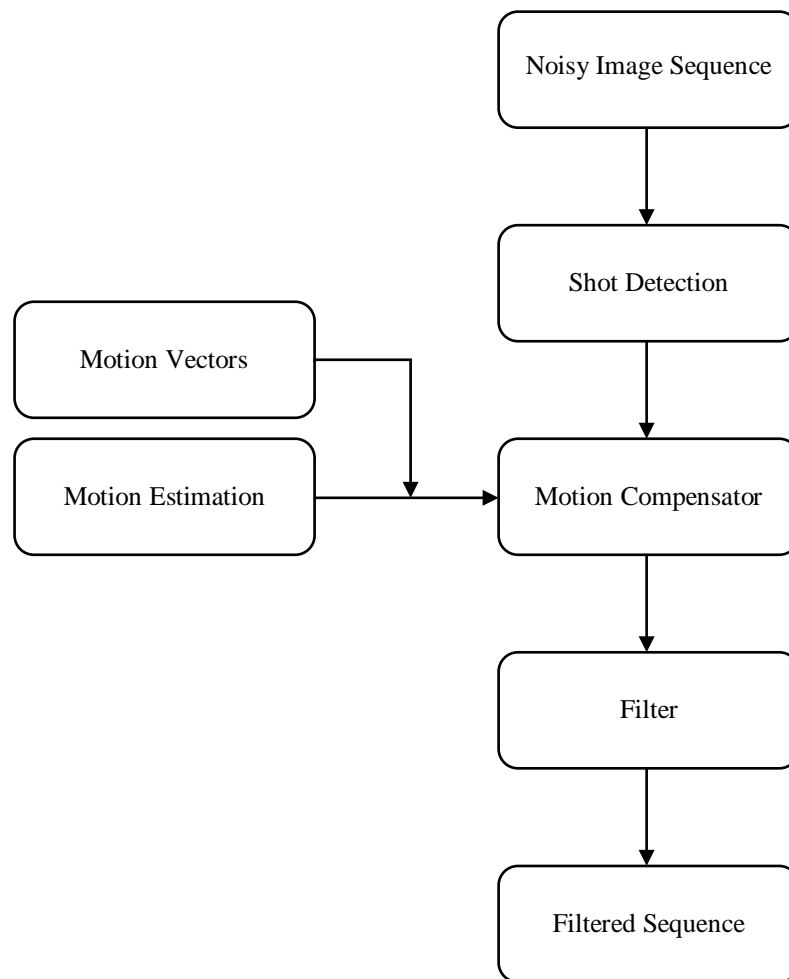


**Figure 7. Deep Learning Model Formation**

### 3. PROPOSED VIDEO DENOISING TECHNIQUE

In the development of the proposed denoising framework for the digital video stream different video stream with different characteristics like frame length and frames per seconds are considered. In the development process, Akio video stream with video length of 300 frames with frame rate of 30 frames per second is considered. The second video stream of football with video length of 260 frames operating on 30 frames per second is considered. The third video input is of soccer with video length of 120 frames with speed of 20 frames per second and the last video stream of traffic monitoring with video length

of 120 frames with frame rate of 15 frames per second is considered. The precise work of the proposed denoising framework is disclosed with the help of the following figure 8.



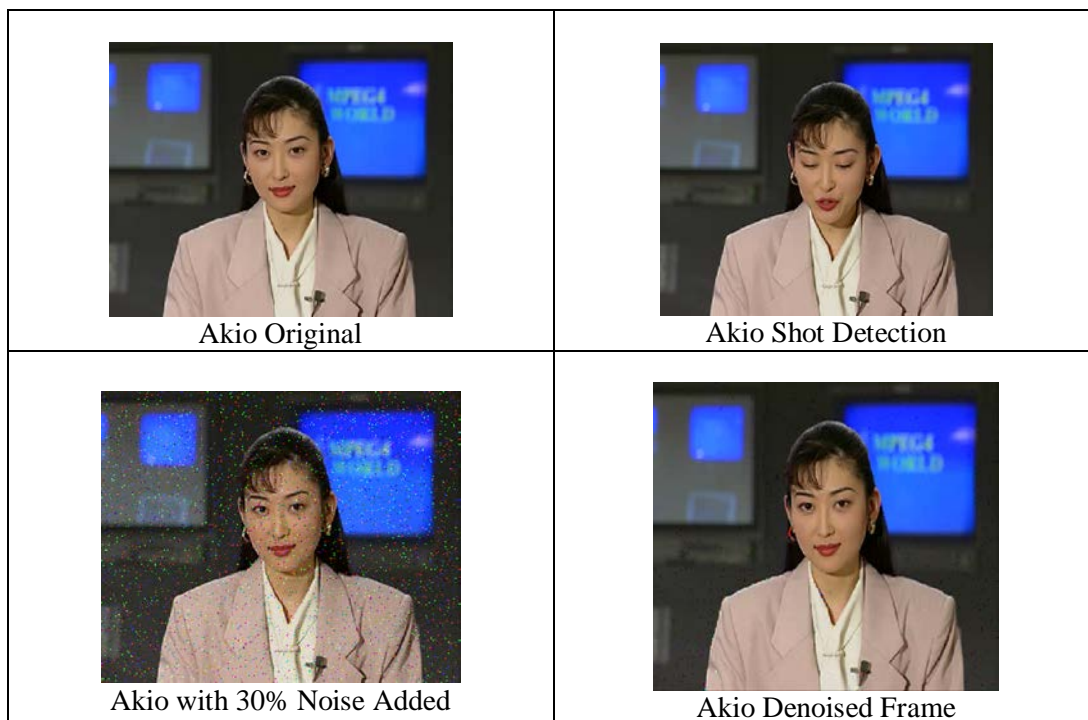
**Figure 8. Work Flow of Video Denoising Framework**

In the next step, the different noise components like Gaussian Noise or Impulse Noise is added in to the video frame(s). Subsequently, the shots are detected in the video frame. In this step segmentation is carried out. It is the step through which only selected components are considered and rest of the unnecessary frame components are deselected from the process. In the next step, motion compensation is carried out. It is the process through which future subsequent frame is estimated. This estimation is carried out using motion vectors and motion estimator. The image intensity at instant (n-1) and at instant (n) is calculated by using noise free frames at instant (n-1), at instant (n) and Gaussian or Impulse Noise sequence at instant (n-1) and at instant (n). Subsequently, motion modelling is done to represent motion of each region, different parameters are initialized and motion for every moving object is estimated and filtering is done to get the final denoised frame at the output. The performance of the proposed denoising framework is evaluated using different ascendancy parameters like Peak Signal to Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM) and Mean Structural Similarity Index Measure (MSSIM). In the subsequent section, the outcome of the proposed framework is disclosed using different input video stream and using statistical values.

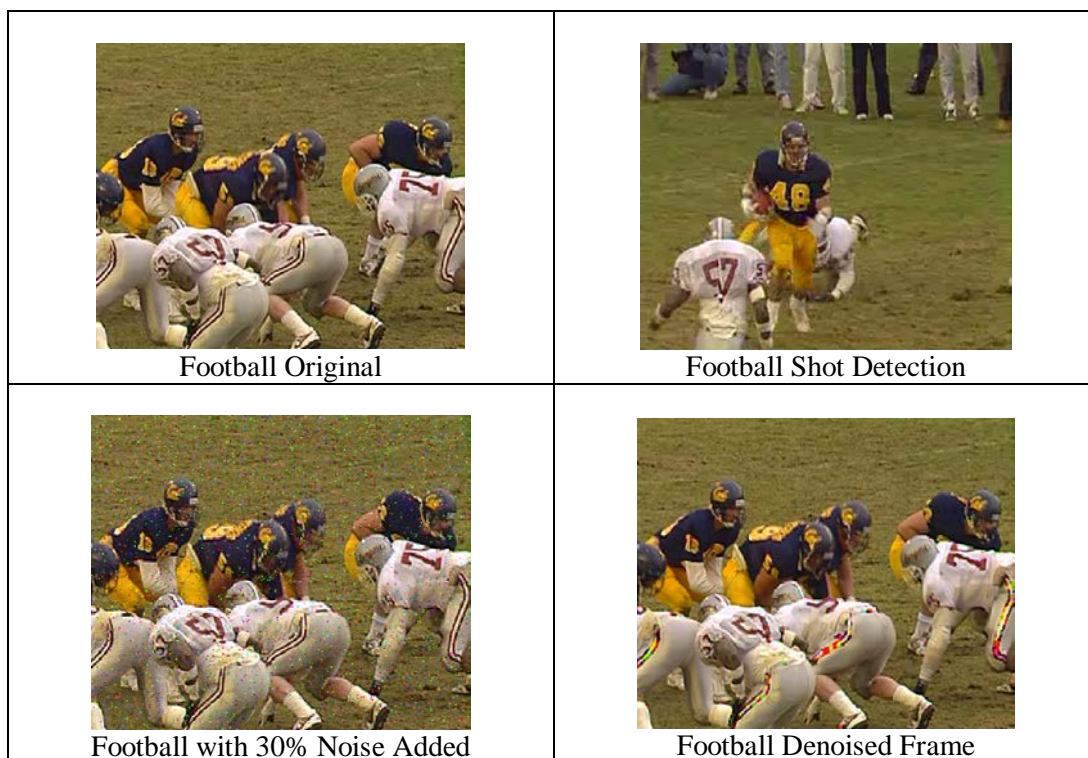


#### 4. RESULT AND DISCUSSION

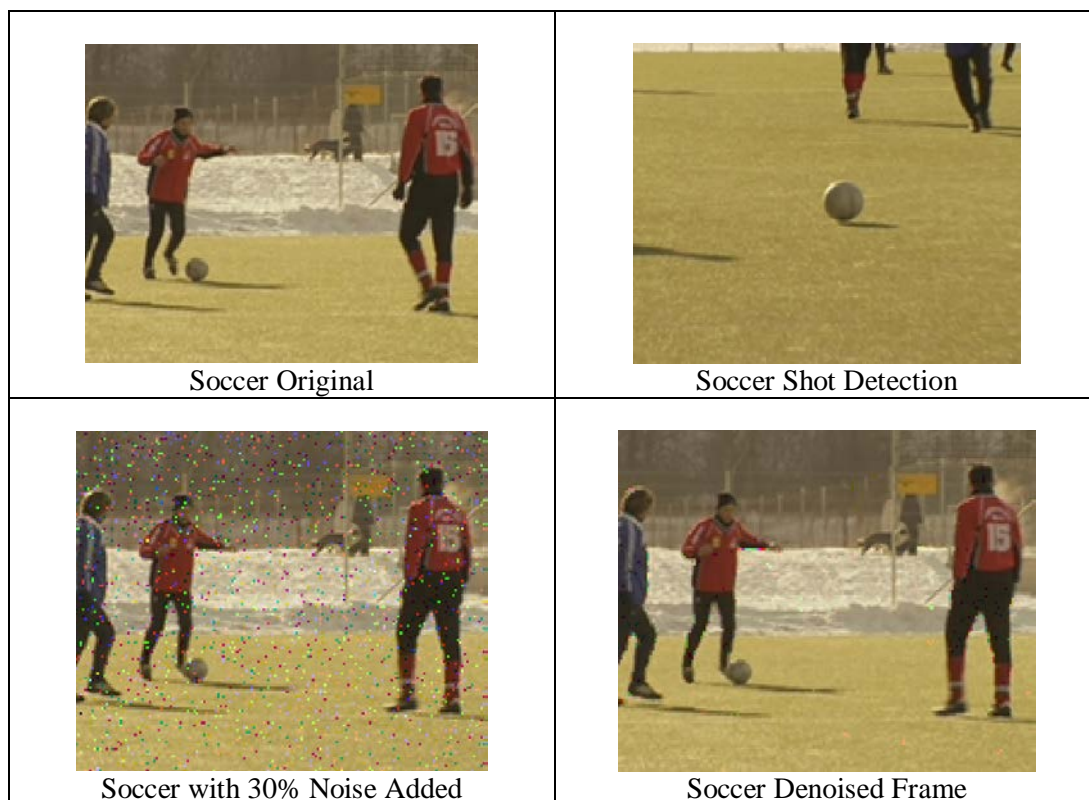
Outcome of the proposed denoising framework for digital video stream is depicted through the images as shown in the figures 9 to 12 below:



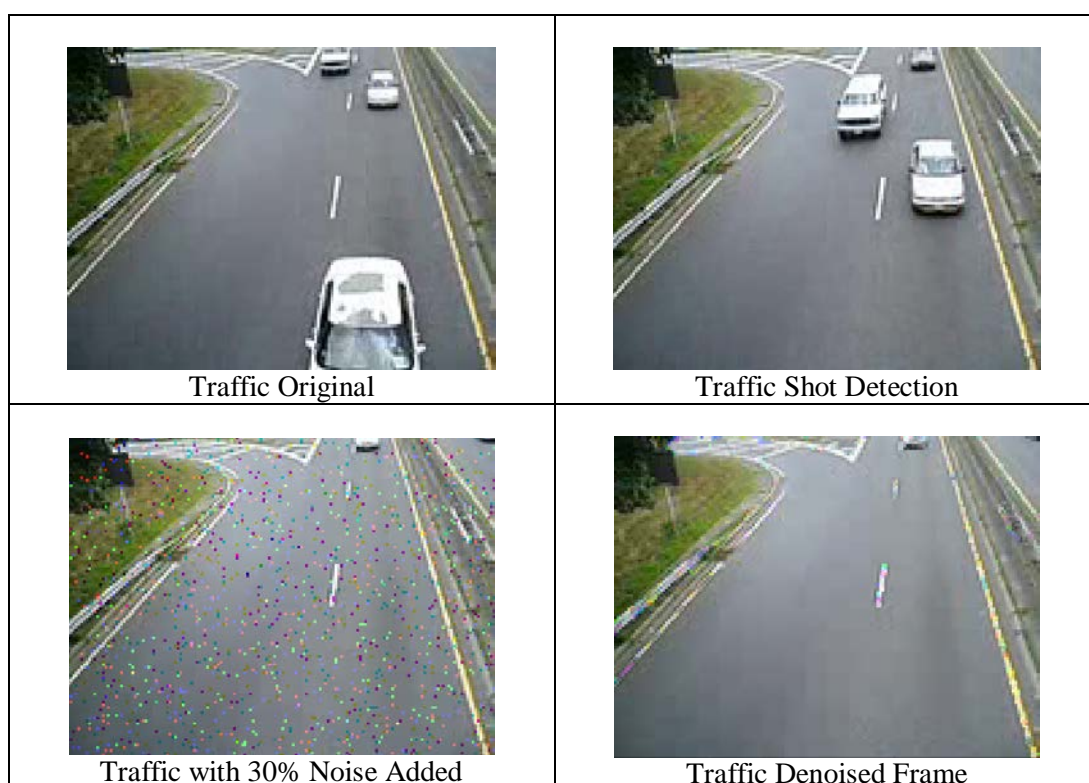
**Figure 9: Effect of Denoising Framework for Raw High Frame Rate Akio Video Stream**



**Figure 10: Effect of Denoising Framework for Raw High Frame Rate Football Video Stream**



**Figure 11: Effect of Denoising Framework for Raw High Frame Rate Soccer Video Stream**



**Figure 12: Effect of Denoising Framework for Raw High Frame Rate Traffic Video Stream**



The statistical analysis of the proposed denoising framework for the different digital video stream is disclosed in the following table 1 to 4:

**Table 1: Statistical Analysis for Akio Video Stream with Video Length = 300 frames and Frame Rate = 30**

% Noise added	PSNR (db)	SSIM	MSSIM
10%	47.1563	0.75242	0.89821
20%	44.0879	0.56916	0.81811
30%	42.4984	0.45576	0.76323
40%	41.2264	0.35773	0.71795
50%	40.2744	0.30268	0.68673
60%	39.4893	0.25668	0.65502
70%	38.8094	0.221	0.63013
80%	38.2608	0.19187	0.60173
90%	37.6623	0.16564	0.58105

**Table 2: Statistical Analysis for Football Video Stream with Video Length = 260 Frames and Frame Rate = 30**

% Noise added	PSNR (db)	SSIM	MSSIM
10%	47.1107	0.83013	0.95682
20%	44.1547	0.71914	0.92706
30%	42.4113	0.68884	0.91132
40%	41.1458	0.55577	0.86868
50%	40.1764	0.49317	0.844
60%	39.2983	0.43091	0.814
70%	38.7131	0.47018	0.8198
80%	38.1419	0.36832	0.77985
90%	37.6558	0.39957	0.77772

**Table 3: Statistical Analysis for Traffic Video Stream with Video Length = 120 Frames and Frame Rate = 15**

% Noise added	PSNR (db)	SSIM
10%	47.3229	0.76323
20%	44.1003	0.57999
30%	42.3975	0.46901
40%	41.1396	0.39329
50%	39.4481	0.30231
60%	38.641	0.25279
70%	38.6431	0.27419
80%	38.0477	0.2329
90%	37.5519	0.209

**Table 4: Statistical Analysis for Soccer Video Stream with Video Length = 120 Frames and Frame Rate = 20**

% Noise added	PSNR (db)	SSIM
10%	47.3229	0.76323
20%	44.1003	0.57999
30%	42.3975	0.46901
40%	41.1396	0.39329

50%	39.4481	0.30231
60%	38.641	0.25279
70%	38.6431	0.27419
80%	38.0477	0.2329
90%	37.5519	0.209

#### 4. CONCLUSION

In this study, we have proposed an effective technique for denoising the video stream. For verification of the proposed framework, we have considered four different video streams each of which are having different video length and different frame rate. In these distinct video streams noise is added in the step of 10% and varied from 10% to 90% value. The performance of the disclosed denoising framework is evaluated using ascendancy parameters like PSNR, SSIM and MSSIM. With the 10% addition of the noise the PSNR value ranges in the house of 47 db, SSIM figures around 0.75 and MMSIM around 0.9. With the increase in the noise value up to 90%, the PSNR drop down to 37 db, SSIM to 0.2 and MSSIM to 0.77.

#### REFERENCES

- [1] <https://www.tubefilter.com/2019/05/07/number-hours-video-uploaded-to-youtube-per-minute/>
- [2] S. Gu, J. Bao and D. Chen, "Learnable Sampling 3D Convolution for Video Enhancement and Action Recognition," 2021 IEEE International Conference on Multimedia and Expo (ICME), 2021, pp. 1-6, doi: 10.1109/ICME51207.2021.9428329.
- [3] M. Li, D. Sun, Y. Lu, Q. Gao and Y. Liu, "Video Image Deblurring Algorithm Based on Denoising Engine," 2021 IEEE International Conference on Information Communication and Software Engineering (ICICSE), 2021, pp. 136-139, doi: 10.1109/ICICSE52190.2021.9404106.
- [4] A. M. Tekalp, M. Covell, R. Timofte and C. Dong, "Editorial: Introduction to the Issue on Deep Learning for Image/Video Restoration and Compression," in IEEE Journal of Selected Topics in Signal Processing, vol. 15, no. 2, pp. 157-161, Feb. 2021, doi: 10.1109/JSTSP.2021.3053364.
- [5] C. Chen, J. Han and Y. Xu, "Video Denoising for the Hierarchical Coding Structure in Video Coding," 2020 Data Compression Conference (DCC), 2020, pp. 362-362, doi: 10.1109/DCC47342.2020.00049.
- [6] J. Enhorn, R. Sjöberg and P. Wennersten, "A Temporal Pre-Filter For Video Coding Based On Bilateral Filtering," 2020 IEEE International Conference on Image Processing (ICIP), 2020, pp. 1161-1165, doi: 10.1109/ICIP40778.2020.9191359.
- [7] H. Yue, C. Cao, L. Liao, R. Chu and J. Yang, "Supervised Raw Video Denoising With a Benchmark Dataset on Dynamic Scenes," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 2298-2307, doi: 10.1109/CVPR42600.2020.00237.
- [8] S. Yu, B. Park, J. Park and J. Jeong, "Joint Learning of Blind Video Denoising and Optical Flow Estimation," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2020, pp. 2099-2108, doi: 10.1109/CVPRW50498.2020.00258.
- [9] A. Li, B. Zheng, L. Li and C. Zhang, "Optical Flow Estimation and Denoising of Video Images Based on Deep Learning Models," in IEEE Access, vol. 8, pp. 144122-144135, 2020, doi: 10.1109/ACCESS.2020.3014345.