Al-Noor Journal of Engineering Management and Computer Science

ISSN: 3079-0689 (Online)

https://njemcs.edu.iq/index.php/njemcs/

Video Denoising Using Convolutional and Deep Neural Networks

Nagwan Badeia Mohamed [1], Waleed Abdullah Araheemah[2] and Ali Abdul Karim Abdullah AL safer[3]

[1] Technical College of Administration - Baghdad Middle Technical University /Baghdad, Iraq.
[2,3]Department of Information Technology Technical -College of Administration - Baghdad / Middle Technical University /Baghdad, Iraq.

ARTICLE INFO

Article history:
Received 24 July 2025
Revised 24 July 2025
Accepted 10 August 2025
Available online 01 September 2025

Keywords:
Gaussian Noise Model (GNM)
Salt-and-Pepper Noise (SPN)
video filtering (VF)
Ideal Low-Pass Filter (ILPF)
Idea High pass filter (HLPF)
Mean Square Error (MSE)

ABSTRACT

Removal of noise from video files has become as one of the most important academic worries and especially with the growing reliance on video data in domains such as security, autonomous vehicles, and entertainment. Recent improvements in deep learning have lead to powerful neural network-based methods that go over traditional Denoising techniques. In this research, a number of deep learning algorithms were tested for video reduced noise on an array of video files. The results demonstrated the effectiveness of deep learning models in reducing noise artifacts, by on the mean square error (MSE) metric, which measures the difference between the original clean videos and the Denoise outputs.

The research included processing (4) video files Which includes a number of videos of different lengths, and the number of frames in this work has become 1,604.through both (convolutional neural network ((CNN)) and (deep neural network (DNN)) and the results showed the superiority of the (CNN) method over the (DNN) method through the results of the (4) deference experiment. The best experiment with minimum MSE was (, μ=0, $\boldsymbol{\sigma}^2$=0.02) with (MSE DNN =0.02001410), the methods can be applied to other multimedia files (audio, image).

1. Introduction

Removing noise from video files is one of the hardest assignments in multimedia processing and analysis. Noise, whether caused by the surroundings or technical limitations during the capture or transmission process, is a significant effect on video quality or affects in the viewer's perception and understanding of the content. Given the increasing reliance on videos in domains such as media, education, and entertainment, enhancing video quality through decreases noise has become essential for delivering a best user experience.

Corresponding author E-mail address: dac2017@mtu.edu.iq
https://doi.org/10.71229/7ajsgv18

Noise in video content can be referred to as unwanted distortions or distortions that reduce visual clarity. It can be due to bad lighting conditions, sensor parameters, analog-to-digital signal conversion errors or even during the compression and storage stages. Traditional denoising techniques, while valuable in certain situations, it often fail to preserve fine visual details or adapt to complex noise patterns.
Recent research trends have shifted toward using deep learning predictive algorithms, particularly Convolutional Neural Networks (CNNs), for video denoising. CNN-based models do automatically learn hierarchical

feature representations that distinguish noise from crucial visual content, enabling more effective noise removal while defending image details.

In this context, this research focuses on leveraging CNN architectures to improve the quality of got or stored photos by effectively removing noise that degrades clarity. The enhancement is particularly vital for applications as television broadcasting, film production, and instruction, where visual quality affects user engagement and comprehension.

For example, in 2022, scientists (Ramesh, G., et al.) studied noise characteristics caused by video transmission and if answers. While the study first employed a Gaussian noise model and Sobel filtering techniques to mitigate noise, new techniques suggest replacing traditional filters with CNN-based models that can adaptively and more accurately reduce noise across diverse video content types, resulting in superior visual restoration performance [1].

Research conducted by (Lindner, Lydia, et al.) in (2023) emphasized the critical value of lowering noise from historical footage, which often stand out by their rarity and historical importance. Enhancing the quality of these archive movies by effective elimination of noise is considered as one of the most significant tasks in multimedia restoration. In their research, the authors highlighted the significance of keeping historical content by removing distortions created essentially by Gaussian noise.

While the original study looked at advanced attention steps, present trends increasingly leverage Deep Neural Networks (DNNs) and Convolutional Neural Networks (CNNs) for similar restoration tasks. CNNs, with their powerful spatial feature extraction capabilities, and DNNs, with their deep hierarchical representations, have demonstrated substantial efficacy in reducing Gaussian noise in historical video files. Recent studies indicates that DNN and CNN-based approaches can significantly enhance the visual quality of rare the past clips by smoothly learning to identify and suppress noise patterns without losing significant image traits. [2]

Devi, B. Aruna, and Mani Deepak Choudhry's (2024) research proposed an entire process for processing remote sensing video files, particularly a focus on noise caused by varying weather conditions during collection of data. The study began by seeing the impact of specialized noise models on the quality of remote sensing captures, with an accent on the challenges associated with environmental distortions.

While the original study employed traditional filtering techniques, such as the median filter based on maximum probability signal distribution models, and compared results using average absolute error metrics, modern developments suggest using Deep Neural Networks (DNNs) and Convolutional Neural Networks (CNNs) for improved results. CNNs may extract spatial features that are critical to determining noise patterns in remote sensing imagery, but DNNs offer robust hierarchical modelling to learn complex noise characteristics related to dynamic weather conditions. Instead of relying totally on handcrafted filters like mean filters or ideal low-pass filters, algorithms for deep learning allow for adaptive, data-driven denoising, leading to improved restoration quality.

In retaining with these improvements, the novel with the extensive availability of video files over the Internet and their frequent transmission across networks, the question of noise contamination in video content has become more important. Noise can significantly diminish video quality, preventing interest levels and seeing rates. As therefore, preserving video clarity has become crucial for ensuring content reach and reach among viewers.

In response to these challenges, deep learning gets closer, particularly those based on Convolutional Neural Networks (CNNs) and Deep Neural Networks (DNNs), have emerged as promising solutions for video denoising. Algorithmic techniques for automate video quality enhancement by learning difficult noise patterns and defining them from essential content factors, strengthening both viewer delight and possibility for closer delivery across digital platforms. The aim of this research shall study and evaluate the performance of various deep learning-based video denoising algorithms, with the focus to enhance the image quality while maintaining key visual qualities inside processed videos. The study seeks to reach a balance between beneficial decrease in noise and retaining of essential factors via using the capabilities of modern neural network forms.

The significance of this research is to develop a greater awareness of the effectiveness of deep learning-based noise removal algorithms applied to video files—an area the fact growing growing in significance as the reliance on digital video builds across an array of industrial, commercial, and daily life applications. Enhancing video quality by effective reducing noise not only boosts visual clarity and user experience, but it also helps to ensure higher dissemination and engagement in digital networks. Leveraging advanced neural network models such as Convolutional Neural Networks (CNNs) and Deep Neural Networks (DNNs), this study strives to create video denoising techniques and their practical use in real-world multimedia events.

strategies use DNN and CNN walls on huge datasets. For example, as processing 50 remote sensing video files with 27 different experimental settings, deep learning models exceeded traditional methods for the purpose of minimizing mean square error (MSE), that lead to a more accurate reconstruction of the original signal while saving critical structural information [3].

Below are some literature reviews related to the topic:

In 2020, researchers Songhyun Yu1 and Bumjun Park1 addressed the problem of video denoising without the use of clean data. To tackle this problem, the researchers stated an interactive learning method that integrates video denoising and optical flow research, in the two neural networks coordinating interactively. The model was tested on the Vimeo-90K database, and performance raised slowly with repeated training. Results showed that the proposed approach outperformed standard techniques, with the network achieving a PSNR of 30.12 dB while denoising video or 28.35 dB when estimating optical flow. The mean square error (MSE) was not explicitly addressed [1].

In 2024, researchers (Zixuan Fu, Lanqing Guo, and Chong Wang) presented an original approach for reducing noise from pictures without a need for clean data. The model was tested on DAVIS, Set8, and CRVD data, and it achieved PSNR = 33.95 dB on Set8 and 35.48 dB in competing with supervised models [14].

In 2025, researchers (Amirhosein Ghasemabadi, Muhammad Kamran Janjua, and Di Niu) devised a way to improve accuracy while reducing resource use. It was tested on data sets including SynNightRain, RVSD, BSD, GoPro, VRDS, DAVIS, Set8, and MVSR4× for rain, snow, blur, noise reduction, and video resolution enhancement tasks. It achieved a PSNR of 34.50 dB for blur removal and 25.30 dB for resolution enhancement, above previous versions. The assessment is based on PSNR and SSIM instead of MSE, showing that it is effective for improving video quality with great efficiency [17].

In 2025, researchers (Kaixun Jiang, Zhaoyu Chen, Jiyuan Fu, Lingyi Hong, and Jinglun Li) addressed the VideoPure model, that improves the security of video recognition against adversarial attacks using Temporal DDIM Inversion to maintain temporal consistency, noise reduction, and Multi-Step Voting to improve prediction accuracy. It was tested on the UCF-101 and Kinetics-400, having defensive accuracies for 85.9% and 80.8%, both (MSE not given) [19].

2. Methodology

2.1. Video Noise

Video noise can be a common problem that affects the quality of videos, especially those shot in low light conditions or using low-quality cameras.

A. Video noise in deep learning-based video processing is a significant issue that requires understanding. Factors contributing to noise growth include low illumination options, sensor limitations in camera hardware, extended exposure time, electromagnetic interference (EMI), lossy compression artifacts, environmental signal interference, and excessive application of image enhancement techniques. Low illumination options require higher ISO sensitivity, leading to significant noise in dark regions. Sensor limitations in low-cost or outdated photon sensors create low signal-to-noise ratios, requiring deep neural networks to compensate for these artifacts. Extended exposure time can induce motion blur and additive noise, making it difficult for models to differentiate between true object boundaries and noise patterns. Electromagnetic interference from nearby devices can corrupt video frames, causing periodic noise that deep learning methods need to adapt to during training.[4].

B. Video noise in video data negatively impacts deep learning performance and visual quality, particularly in tasks like classification, object recognition, motion estimation, and video compression. It impairs visual quality by introducing random or structured distortions that degrade frame spatial coherence, suppresses fine-grained features, increases file size in enhanced videos, induces visual fatigue in human perception, and lowers compression efficiency due to increased entropy in video frames. However, video noise removal can provide substantial advantages in both human perception and machine analysis domains. Deep learning-based denoising methods enhance video clarity and sharpness by reducing noise artifacts, improving detail resolution, and enhancing the human viewing experience. Noise-free content also eases cognitive stress and eye strain, making it essential in entertainment media, virtual systems, and content delivery platforms. Higher compression efficiency is achieved by removing noise, allowing modern codecs to achieve high compression ratios with reduced perceptual quality loss. Clean input frames lead to more accurate feature extraction, boosting object detection, semantic segmentation, activity recognition, and motion tracking performance.

   Video denoising is a crucial part of post-production workflows in film, screens, and advertising industries, as high-fidelity, noise-free visuals reflect professionalism and elevate content credibility. Integrating reduced sound into deep learning improves not only the human sense of quality but also model generalizability, training stability, and downstream performance in both supervised and unsupervised video analysis tasks.

C. Video quality in AI-powered video enhancement systems is influenced by various factors, including video resolution, which determines the number of pixels per frame and affects detail for human perception and model input. Higher resolutions improve feature extraction but also have more computational and storage capacity during training and inference. Temporal resolution, measured in frames per second, affects motion continuity and smoothness, with videos with higher clip rates providing more stable data, requiring motion-aware models like ConvLSTMs and spatiotemporal transformers.

D. Video quality in AI-driven video processing is influenced by various technical and environmental factors. The most critical factors include resolution, which defines the number of pixels in each image, which affects the level of detail in visual content. Higher resolutions provide more input for deep learning models but require more computational resources and memory. Frame rate, measured in FPS, controls how smoothly motion is displayed, enhancing temporal coherence for deep models like ConvLSTM or 3D-CNNs. Bitrate, an amount of video data sent per second, preserves more visual information but increases file size.

E. Video quality in deep learning-based video processing is influenced by various factors, including video resolution, frame rate, bit rate, and lighting conditions. High video resolutions provide richer spatial detail, but increase computational complexity and memory use. Frame rate affects the temporal continuity of motion information, benefiting CNNs with temporal extensions. High-bitrate content preserves the original structure, improving model robustness and output quality. Lighting conditions also play a role, as poor lighting increases noise rates, especially in darker areas. These factors affect the performance of CNNs and DNNs in video processing.
F. Camera models and sensor quality play a crucial role in the learning process of CNNs and DNNs. Advanced sensors with better dynamic range and light sensitivity provide cleaner frames, while low-end sensors may require more pre-processing or model regularization to minimize noise. Lossy compression methods can induce structural artifacts, requiring deep networks to discriminate between true content and artifacts. Videos with greater bit-depth offer better color information, useful for CNNs in tasks like color restoration or super-resolution. Post-processing effects, such as color correction and noise reduction, can impact the distribution of pixel values, requiring DNN models to generalize to both raw and post-processed data. Noise characteristics, such as uniform, Gaussian, Salt & Pepper, and Rayleigh noise types, also affect video quality.

2.2. Video Models

Noise models in video link to the many different kinds of visual distortions that may appear in video frames, either during recording or during processing. Understanding these models is essential in designing effective algorithms to reduce or eliminate noise to enhance video quality.

Artificial noise is often added to videos in research and experimentation to evaluate the value of denoising methods or to train deep learning models for imitating the original clean frames. The sort of noise built varies based on the conditions being reused, such as poor light. Compression issues or the usage of low-quality imaging tools:

A. Gaussian Noise Model(GNM)

Gaussian noise is a common type of noise employed in video evaluation and testing. It is also known as Gaussian white noise and is marked by a random distribution with a constant variance and zero mean.

When used on a video, Gaussian noise appears as fine grainy distortions—small dots or specks distributed in the frames. It typically impacts darker and mid-tone regions, cautiously lower visual quality.

Noise at pixel coordinates (x,) can often be defined using a Gaussian (normal) probability distribution function with mean μ and variance σ, stated as:
4]

$$P(n) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(n-\mu)^2}{2\sigma^2}} \qquad (1)$$

- In real-world situations, Gaussian noise is added to video frames to simulate scenarios like low light, electronic sensor interference, or transmission defects. The controlled inserting of noise is crucial in determining the effectiveness and effectiveness of denoising algorithms in real-life situations.

B. Salt-and-Pepper Noise(SPN)

Salt-and-pepper noise is frequently employed in computer vision and deep learning studies to simulate simple distortions and test denoising applications' toughness. This sort of noise appears as a mix of black (pepper) and white (salt) pixels across video frames, greatly lowering visual quality and obscuring tiny details.

Such noise typically occurs from digital

transmission errors, hardware problems in image sensors, or quantization mistakes. Salt-and-pepper noise has been deliberately added to deep learning models, primarily convolutional neural networks (CNNs), so that the model learn to identify and remedy unforeseen intensity interruptions.:[5]
A probabilistic model generates the intensity of a noisy pixel (f) at location (i, j).:

$$f(\acute{i,j}) = \{ \begin{matrix} 0 & p(\text{pepper}) \\ f(i,j) & \left(1 - p(\text{pepper}) + p(\text{salt})\right) \\ 255 & p(\text{salt}) \end{matrix} \quad (2)$$

2.3. Video Filtering

Video denoising is the process of removing unwanted noise from video frames to enhance visual quality and restore lost attributes. This goal is achieved via a selection of methods, including traditional filtering methods and electricity deep learning methods.

The Low-Pass Filter is a typical traditional method to decreasing high-frequency components which are frequently associated with noise. This filter smooths pixel values by averaging them with their neighbours, efficiently reducing sudden intensity shifts and preserving low-frequency news like object methods and soft textures.
.[6].

A .low pass l filtering

Spatial filters are often used in video processing to enhance video quality by altering pixel values based on the hue of neighbouring pixels within the same frame. These filters operate directly in the spatial domain, making them helpful in video smoothing, contrast enhancement, and edge recognizing. The Mean Filter is a simple widely used spatial filter. It reduces noise by replacing the intensity of each pixel with the average of the pixels near it. The result well prevents tiny, random noise while giving a smoothing effect over the frame. However, it may result in the loss of fine details, in particular for areas that includes data at high rates such as edges and textures.

The mean filtering operation can be written by numbers as
As: [7]-

$$g(i,j) = \frac{1}{m * n} \sum_{m=-a}^{a} \sum_{n=-b}^{b} f(i+m, j+n) \quad (3)$$

Such that
$(g(i,j) is\ the\ output\ pixel\ value$
f (i,j) is the original input pixel value
M*N is the size the filter window (e.g , 3*3)
a = $\frac{M-1}{2}, b = \frac{N-1}{2} define\ the\ window\ bounds.$
($f(i+m, j+n)$) represent the pixel value.

B. High pass filtering

Frequency domain filtering is a video and signal processing technique that alters a video's frequency components rather than the spatial pixel values. These filters may be referred to as frequency filters as they target specific frequency bands to enhance or suppress particular features in the signal.

One of the most common uses for frequency filters is noise reduction, especially within high-frequency regions. They are also widely used in edge enhancement and information extraction, where frequency components give an improved grasp of the structural characteristics of visual video. A famous instance is the Ideal Low-Pass Filter (ILPF), which allows low-frequency components to pass while completely attenuating high-frequency components. This filter is essential for video smoothing, as it reduces sudden shifts and suppresses high-frequency noise. The ILPF is conceptually simple and good for softening or blurring signals, yet it can eliminate some high-frequency traits that are important to texture and edge identity.[8]
.

$$H(u,v) = \{ \begin{matrix} 1 & if\ \ D(u,v) \le D_0 \\ 0 & if\ \ D(u,v) > D_0 \end{matrix} \quad (4)$$

$$D(u,v) = \sqrt{(u - \frac{M}{2})^2 + (v - \frac{N}{2})^2} \quad (5)$$

(D_0) cutoff frequency

$$F(u,v) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} f(x,y) e^{-2\pi j [\frac{ux}{M} + \frac{vy}{N}]} \quad (6)$$

$$g(x,y) = \frac{\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} F(u,v)e^{2\pi j[\frac{ux}{M}+\frac{vy}{N}]}}{MN} \qquad (7)$$

3. VIDEO DENOISING ALGORITHMS

Removing noise from video is an important process in improving the quality of the video and making it clearer, especially in applications such as surveillance, movies, medical images, etc. There are several algorithms foremoving noise from video based on different techniques for processing frames and reducing noise.[8]

A.Convolutional Neural Network-Based Video Denoising Algorithm (CNN-VDA)

CNN-based video denoising strategies shorten noise from each video frame using deep learning models than were provided to distinguish between clean and noisy content. CNNs use convolutional layers to automatically learn hierarchical features from data, as opposed to traditional spatial domain arrives at that rely on fixed pixel-wise methods. These algorithms rate each frame independently, using convolutional filters to find local and global patterns in the image. By training the network on huge sets of noisy and clean frames, the CNN learns to reconstruct Denoise versions that retain important visual detail while well suppressing different kinds of CNN-based nears are particularly effective in removing planned and complex noise patterns, such as Gaussian or Salt-and-Pepper noise, and coping with an array of video footage. Basic CNNs, as spatial gets closer, do not take into account time correlations between consecutive frames unless they are raised to 3D CNNs or recurrent structures. [9]
It can be represented by the following algorithm figure 1.

Algorithm Steps
Input video file
Add Gaussian noise with (μ, σ^2)& Salt-and-Pepper Noise with (p)
Apply denoising using: low –pass filter, high –pass filter, cnn algorithm, and dnn algorithm
Calculate the mean square error (mse)

Figure 1. Shows proposed video denoising and evaluation Algorithm.

B. *Deep Neural Network Video Denoising Algorithm (DNN-VDA)*

DNN-based video denoising modes create refined mappings between noisy and clean video frames via deep neural networks. Instead of using standard frequency domain alterations such as the Fourier Transform, such techniques use many fully connected or convolutional layers to find noise characteristics and create cleaner frames.
It starts with feeding noisy video frames into a neural network, which has been trained on big datasets with matched samples of noisy and clean frames. Through backpropagation and optimization approaches, the DNN learns to lessen the reconstruction error—commonly defined as Mean Squared Error (MSE)—by modifying its weights to suppress noise while preserving vital visual details. [10].

4. VIDEO DATA BASE

The research database included a number of video files) four (according to different lengths. Figure 2 represents one of these files.

Figure 2. One of Video data base

5. SUGESSTED METHODOLGY

The methodology for applying the two algorithms to a number of video files according to figure 3.

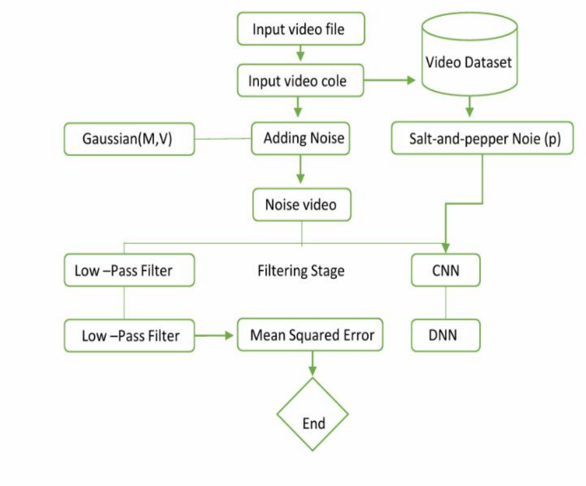

Figure 3. methodology research

5. NUMERICAL RESULTS

Results are represented in the following tables and figures:

Table1: Gaussian noise results main=0.sigma=0.02 by CNN algorithm:

Video	MSE Noisy	MSE Low Pass	MSE High Pass	MSE CNN
Avi.avi	0.00040187	0.00080236	0.3592	0.026567
Mov.mov	0.0003997	0.0038955	0.23973	0.03113910
Mp4.mp4	0.00040071	0.0033727	0.16399	0.0227717
Mpeg.mpeg	0.00039891	0.0023846	0.17815	0.0174417

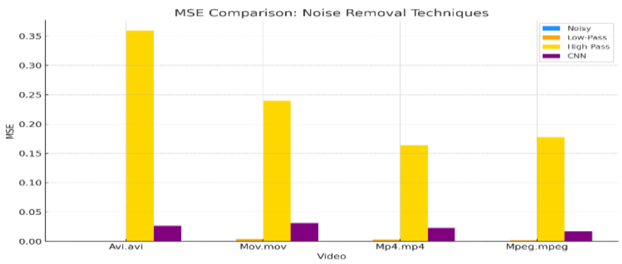

Figure 4. Results of Gaussian noise main=0.sigma=0.02 by CNN algorithm

Table 2: Salt and pepper noise results when p1=0.01 by CNN algorithm

Video	MSE Noisy	MSE Low Pass	MSE High Pass	MSE CNN
Avi.avi	0.0030034	0.0010444	0.3614	0.02532310
Mov.mov	0.0029247	0.0041688	0.23942	0.02994710
Mp4.mp4	0.0332961	0.00368	0.16491	0.02201317
Mpeg.mpeg	0.0035095	0.0026977	0.1796	0.0174437

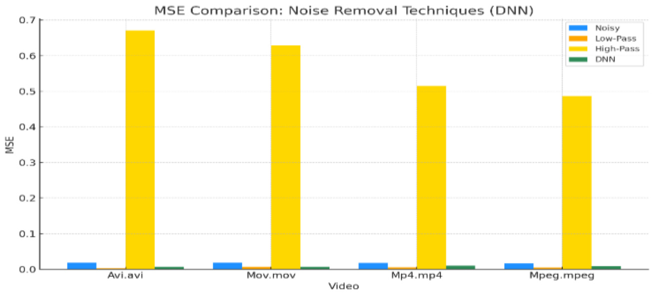

Figure 5. Salt and pepper noise results when p1=0.01 by CNN algorithm

Table 3: Gaussian noise results main=0.sigma=0.02 by DNN algorithm

video	MSE Noisy	MSE Low Pass	MSE High Pass	MSE DNN
Avi.avi	0.018249	0.0028586	0.67037	0.007182210
Mov.mov	0.019246	0.0064553	0.6293	0.006892910
Mp4.mp4	0.017798	0.0059674	0.51496	0.00991817
Mpeg.mpeg	0.017081	0.0052917	0.48602	0.00856027

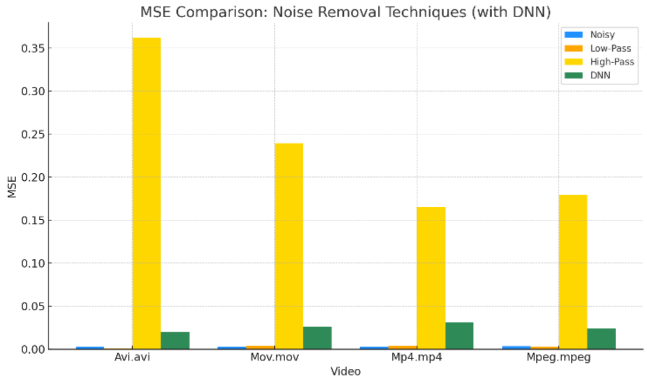

Figure 6. Salt and pepper noise results when p1=0.02 by CNN algorithm

Table 4: Salt and pepper noise results when p1=0.01 by DNN Algorithm

video	MSE Noisy	MSE Low Pass	MSE High Pass	MSE DNN
Avi.avi	0.0027604	0.0010176	0.36163	0.02001410
Mov.mov	0.0028751	0.0041664	0.23933	0.02635310
Mp4.mp4	0.0032693	0.0036699	0.16509	0.03121617
Mpeg.mpeg	0.0034709	0.0026805	0.17924	0.0244897

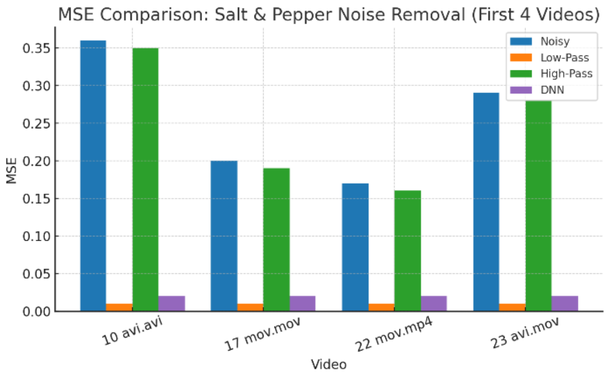

Figure 7. Salt and pepper noise results when p1=0.01 by CNN algorithm Database

5. Conclusions and Suggestions

Several major inferences have been drawn after analysing results for each experiment that had been carried out. The convolutional neural network (CNN) surpassed the deep neural network (DNN) in the evaluation of deep learning methods of noise reduction, with an MSE of 0.0174437 as compared to 0.02001410 for the DNN. These findings show CNN-based filtering is interest for helping preserve video quality after the reduction of noise. Table III provides a thorough overview of prior study, offering insights into the creation and use of image classification technique in that field.

References

[1] Davy, Axel, et al. "A non-local CNN for video denoising." *2019 IEEE international conference on image processing (ICIP)*. IEEE, 2019.

[2] Ho, Man M., Jinjia Zhou, and Gang He. "RR-DnCNN v2. 0: enhanced restoration-reconstruction deep neural network for down-sampling-based video coding." *IEEE Transactions on Image Processing* 30 (2021): 1702-1715.

[3] Liu, Peng, et al. "Remote-sensing image denoising with multi-sourced information." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 12.2 (2019): 660-674.

[4] Rashid, Loay, Siddharth Roheda, and Amit Unde. "LLVD: LSTM-based Explicit Motion Modeling in Latent Space for Blind Video Denoising." *arXiv preprint arXiv:2501.05744* (2025).

[5] Gonzalez, R. C., & Woods, R. E. (2018). *Digital Image Processing* (4th ed.). Pearson

[6] **Xu, J., Zhang, L., & Zuo, W. (2022).** *Noisy Image Denoising Using Deep CNN With Batch Normalization and Skip Connections. Signal Processing: Image Communication, 97, 116548.* https://doi.org/10.1016/j.image.2021.116548

[7] Birkfellner, Wolfgang. Applied medical image processing: a basic course. CRC Press, 2024.

[8] Gonzalez, R. C., & Woods, R. E. (2018). *Digital Image Processing* (4th ed.). Pearson..

[9] Tassano, M., Delon, J., & Veit, T. (2019). **FastDVDnet: Towards Real-Time Deep Video Denoising Without Flow Estimation.** *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).*

[10] Tassano, M., Delon, J., & Veit, T. (2020). **FastDVDnet: Towards Real-Time Deep Video Denoising Without Flow Estimation.** *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020*

[11] Yoon, Sunjae, et al. "FRAG: Frequency Adapting Group for Diffusion Video Editing." arXiv preprint arXiv:2406.06044 (2024).