ENHANCING VIDEO QUALITY THROUGH AI-POWERED NOISE REMOVAL TECHNIQUES

Project Reference No.: *47S_BE_3405*

College ***:*** *Shri Madhwa Vadiraja Institute of Technology and Management, Udupi*
Branch ***:*** *Department of Computer Science and Engineering*
Guide(s) ***:*** *Ms. Sahana*
Student(S) ***:*** *Mr. Krishnamoorthy Karantha K.*
Ms. Arathi Kini
Ms. K. Prathvi Rao
Mr. Nandeesh U.

Keywords:
Video Denoising, Noise Reduction, Gaussian Blur, Bilateral Filter, Non-Local Means, Deep learning, DnCNN, CLSTM, U-Net Autoencoder, PSNR, MAE, SSIM.

Introduction

High-quality video content is crucial in today's digital age for engaging audiences across various platforms. However, achieving pristine video quality is challenging due to noise and artifacts from low-light conditions, compression, data transfer errors, and sensor limitations. Noise manifests as random variations in brightness or color, degrading visual clarity and detail. Traditional denoising techniques like Gaussian Blur and Fast Non-Local Means help but often fail to balance noise removal with detail preservation.

Noise in videos is classified into temporal noise (fluctuating lighting), spatial noise (sensor irregularities), and compression artifacts (encoding/decoding errors). Effective denoising must address these categories, ensuring computational efficiency for real-time applications like live streaming, robustness to various noise levels and types, adaptability to different content, and preservation of visual features.

This study compares state-of-the-art AI-based denoising models: Convolutional Long Short-Term Memory (CLSTM), Denoising Convolutional Neural Network (DnCNN), and Autoencoder models. Using diverse video datasets and simulating real-world noise scenarios, we evaluate each model's performance through quantitative metrics (PSNR, processing time) and qualitative visual assessments.

Additionally, we examine computational efficiency and scalability to determine practical applicability. This comprehensive analysis aims to identify superior denoising solutions leveraging deep learning architectures.

Objectives:

- To enhance video content by reducing noise, artifacts, and maintaining clarity, delivering cleaner and more professional content.
- To compare different AI video denoising models

Methodology:

. Materials

a. Dataset:

1. The DAVIS 2017 dataset is used, including high-quality video sequences with pixel-wise annotations for object boundaries.

2. This dataset comprises 50 video sequences, split into a training set (80%) with 40 video sequences and a testing set (20%) with 10 video sequences.

b. Software and Libraries:

1. Python is used for implementing and evaluating the denoising techniques.

2. OpenCV is utilized for traditional video denoising techniques.

3. TensorFlow is employed for implementing the DnCNN, CLSTM, and U-Net Autoencoder models.

c. Hardware:

1. GPU/TPU processing power is essential for efficient model training.

2. Adequate VRAM and SSD storage are critical for handling large video datasets.

B. Methods

a. Data Collection and Preprocessing:

1. The DAVIS 2017 dataset is collected and split into training and testing sets.

2. Video frames are normalized, resized to 256x256 pixels, and converted to the appropriate color

space.

b. Model Selection and Implementation:

1. Traditional Techniques:

2. Gaussian Blur smooths variations, reducing noise but potentially blurring details.

3. Bilateral Filter minimizes noise while preserving edges by considering spatial proximity and intensity similarity.

4. Fast Non-Local Means (NL Means) Denoising uses weighted averages to remove noise and preserve details.

5. Deep Learning Models:

6. DnCNN uses convolutional layers, batch normalization, and ReLU activations to map noisy images to clean outputs.

7. CLSTM combines convolutional layers with LSTM units to capture spatial and temporal features.

8. U-Net Autoencoder uses an encoder-decoder network with skip connections to capture features for effective denoising.

c. Model Training:

1. DnCNN: Processes 256x256 RGB images using stacked Conv2D layers for feature extraction and image reconstruction.

2. CLSTM: Uses ConvLSTM2D layers for video sequences, focusing on spatial and temporal features with BatchNormalization.

3. U-Net Autoencoder: Employs convolutional layers in both encoder and decoder stages with dropout to prevent overfitting.

d. Hyperparameter Tuning:

1. The Adam optimizer is chosen for its adaptive learning rates.

2. The learning rate is set to 0.01 for balanced convergence speed and stability.

3. Mean Squared Error (MSE) is used as the loss function.

4. Early stopping with a patience of 2 epochs is applied to prevent overfitting.

e. Model Evaluation:

1. Metrics:

2. Processing time per frame is measured in seconds.

3. Peak Signal-to-Noise Ratio (PSNR) measures the ratio between maximum signal power and noise power.

4. Mean Absolute Error (MAE) calculates the average absolute difference between original and denoised frames.

A. Evaluation of Traditional Denoising Techniques:

Our evaluation compared three noise removal techniques: Gaussian blur, bilateral filter, and Non-Local Means (NL Means) denoising, using four metrics: processing time, Peak Signal-to-Noise Ratio (PSNR), Mean Absolute Error (MAE), and Structural Similarity Index Measures (SSIM).

i. Processing Time: Gaussian blur showed the fastest processing time, but it may blur details in the video due to its simple operation.

ii. PSNR: Gaussian blur may increase PSNR by reducing noise, but it can sacrifice fine details, whereas NL Means achieved a better balance.

iii. MAE: NL Means typically yielded lower MAE values, suggesting better preservation of image details while reducing noise.

iv. SSIM: NL Means often achieved higher SSIM scores, indicating better preservation of overall structural information.

Key Findings: Gaussian blur prioritizes speed over detail preservation, while NL Means offers a better balance between noise reduction and detail preservation.

B. Evaluation of Deep Learning Models:

Our evaluation compared DnCNN, CLSTM, and U-Net Autoencoder in terms of processing time and PSNR.

i. Processing Time: CLSTM exhibited quicker processing times due to its simpler architecture, while DnCNN showed superior PSNR.
ii. PSNR: DnCNN outperformed CLSTM and U-Net AutoEncoder in PSNR, prioritizing signal fidelity over processing speed for better noise reduction

performance.

Key Findings: CLSTM prioritizes processing speed, potentially at the expense of detail preservation, while DnCNN excels in noise removal quality.

C. Conclusion:

This study explored various video denoising techniques, traditional and deep learning-based, revealing a trade-off between processing speed and detail preservation. Gaussian blur prioritizes processing speed but may compromise details preservation. NL Means offers a more balance approach, achieving noise reduction while maintaining image details to a greater extent. CLSTM stands out for its emphasis on processing speed. DnCNN excels in noise removal quality, providing a superior balance between speed and denoising effectiveness.

Description Of the Innovation in The Project

DEEP LEARNING PROJECT
The project utilizes deep learning techniques to solve complex problems by training neural networks with large datasets. Deep learning, a subset of AI, enables more accurate predictions and improved efficiency across various domains. This innovation extends to tasks like image recognition, natural language processing, and predictive analytics, pushing the boundaries of technology. Deep learning models have demonstrated remarkable success in numerous applications, from healthcare to finance and autonomous driving. Overall, the project's innovation lies in harnessing deep learning's capabilities to solve real-world problems and drive advancements.

Future Work Scope:

While our research has shown promise, further investigation is needed to identify the optimal deep learning model for video denoising. We plan to explore additional evaluation metrics like Mean Absolute Error (MAE) and Structural Similarity Index Measure (SSIM) for a comprehensive assessment of model performance. Ongoing efforts will focus on evaluating deep learning models more extensively to achieve superior noise reduction capabilities. Additionally, we aim to enhance processing speed by utilizing parallel processing techniques and leveraging GPU acceleration. Integrating these findings with our current research will provide a more thorough understanding of video quality enhancement methods.