IMAGE DENOISING USING RESOFOCUS AND FRAGMENTUMZOOM AUTOENCODERS

[1]B.Poojitha, [2]V.Subhash, [3]P.Harshithsai, [4]U.Nandha Kishore, [5]P.Yashwanth Kumar, [6]P.V.R.Srinivas Kumar

[1]Assistant Professor, [1]Assistant Professor, [2,3,4,5]UG Student, [1,2,3,4,5,6]Department of Computer Science & Engineering (AI&ML), Geethanjali Institute of Science And Technology, Nellore, India

Abstract

Image denoising is the process of removing noise from images to make them clearer and more useful. Auto encoders, a type of neural network, are widely used for this task because they can learn to compress and reconstruct images effectively. Recent innovations in auto encoder architectures, such as adding attention mechanisms or multi-scale processing, have made denoising even better. These advancements help preserve important details while removing unwanted noise, making them useful in fields like medical imaging, photography, and satellite imaging. This work explores how auto encoders can be improved to achieve high-quality image denoising in a simple and efficient way and here using the efficient algorithms like Reso Focus auto encoder and Fragmentum Zoom Architectures.

Keywords: Image denoising, Reso focus, fragmentumzoom, auto encoders

Introduction

In the digital era, high-quality visual data is essential for a wide range of applications such as photography, satellite imaging, autonomous driving, medical diagnostics, and surveillance systems. Despite advancements in imaging technologies, images often sufferfrom noise due to hardware limitations, environmental factors, or transmission errors. Noise, characterized by random pixel intensity variations, can significantly degrade image quality, obscure important features, and hinder downstream tasks like object detection and classification. Image denoising, therefore, is a fundamental challenge in computer vision, aimed at restoring a clean image while preserving crucial elements such as edges, textures,and structures.

This project proposes a novel, data-driven approach to image denoising using twodeep learning architectures — Fragmentum Zoom Autoencoder and ResoFocus Autoencoder. These models address the shortcomings of traditional filtering methods and standard deep learning techniques, which often blur fine details. The FragmentumZoom Autoencoder employs a fragment-based, zoom-level learning approach, dividing images into sections and processing them at multiple zoomscales. This enablesthe modelto focuson fine texturesand localized patterns that are especially vulnerable to noise. In contrast, the ResoFocus Autoencoder uses multi-resolution and residual learning to capture both global structures and local details, effectively enhancing clarity while avoiding over-smoothing.

Together, these architectures offer a powerful solution for denoising under various noise types, including Gaussian, salt-and-pepper, and speckle noise. Evaluations show that both models outperform conventional techniques in terms of quantitative metrics like PSNR and SSIM, aswell as visualquality. Thisdual-modelframework not only advances intelligent image restoration but also illustrates the potential of custom neural network designs intackling specific data corruption challenges. The techniques presented can be extended to domain-specific tasks such as low-light image enhancement, satellite image denoising, and medical scan restoration, demonstrating the versatility and impact of the proposed approach.

Image Processing

Image processing is the use of computers to enhance, modify, or analyze digital images.

Noise

Noise in images can be a quite a problem. It often occurs when you record an image (e.g. take a photo) without having enough light available. We can think of noise as adding random numbers (positive and negative) to each pixel value. So, in a noisy image some pixel values will be brighter than they should be and some will be darker.

Types of Noise

Gaussian Noise

Randomvariationsinpixelvaluesthatlooklikegrainyspeckles,oftencausedbysensor limitations.

Salt-and-Pepper Noise

Appearsasrandomblackandwhitepixels,usuallyduetodatacorruptionortransmission errors.

Poisson Noise (Shot Noise)

Causedbyvariationsinlightlevels,commoninlow-lightormedicalimaging.

Image Denoising

Imagedenoisingistheprocessofremovingunwantednoiseordistortionfromadigitalimage to make it clearer.

Image Denoising Techniques

Spatial Filtering

Median Filtering

Replaces eachpixel's value withthe median value ofits neighboring pixels, effectively removing impulse noise and preserving edges.

Others patial filters

Include Gaussian filters, which smooth images by averaging pixelvalues within a neighborhood, and bilateralfilters, which preserve edges while smoothing noise.

Transform Domain Filtering

Wavelet Denoising

Transforms the image into a wavelet domain, where noise is often concentrated in high-frequency coefficients. Thresholding these coefficients removes noise while preserving image details.

Fourier Transform

Transforms the image into the frequency domain, allowing for filtering of specific frequencies that correspond to noise.

Machine Learning/Deep Learning

Convolutional Neural Networks (CNNs)

Trained on pairs of noisy and clean images, CNNs learn to map noisy inputs to clean outputs.

Generative Adversarial Networks (GANs)

Employ two networks, a generator that produces denoised images and a discriminator that evaluates their quality, leading to better denoising results.

Deep Learning

Deep learning offers a powerful approach to image denoising byleveraging convolutional neural networks (CNNs) and other architectures to learn and remove noise from images.These models learn to map noisy images to their corresponding clean versions, effectively reducing noise artifacts.

Deep Learning-Based Image Denoising Techniques

Convolutional Neural Networks(CNNs)

CNNs are widely used for image denoising due to their ability to learn hierarchical features from image data. They can directly map noisy images to their corresponding clean versions, trained to minimize the difference between the denoised output and the ground truth.

Denoising Auto encoders (DAEs)

DAEs are a type of neural network that learns to reconstruct the input image from a noisy version. They are trained to remove noise by forcing the network to prioritize the most important features of the original image.

Generative Adversarial Networks(GANs)

GANs, particularly those with conditional GANs, can be used to learn the underlyingdistributionof cleanimagesandthengeneratedenoisedimagesfrom noisy inputs. They involve a generator network that attempts to create realistic clean images and a discriminator network that tries to distinguish between the generator's output and actual clean images.

Hybrid Approaches:

Combining different deep learning models or techniques, such as CNNs with generative models or DAEs with other denoising methods, can improve performance and overcome the limitations of individual approaches.

Multi-Stage Networks:

Some deep learning models use multiple stages to progressively remove noise and refine features, improving the overall denoising quality and handling complex noise types.

Auto Encoder

An auto encoder is an algorithm designed to generate an output image closely resembling the input image, and its true utility becomes apparent with a more insightful definition. Essentially, an auto encoder is crafted to learn an informative representation of data through the precise reconstruction of a set of input observations, enabling versatile applications

Architectural Components

To comprehend auto encoders, it is essential to explore their typical architecture, The fundamental components of autoencoders encompass an encoder, a latent feature representation, and a decoder. The encoder and decoder operate as functions complemented by the latent feature representation, typically denoting a tensor of real numbers. The overarching objective is twofold: ensuring effective reconstruction of input data while concurrently creating a meaningful latent representation.

Encoder

In image denoising using autoencoders, the encoder component plays a crucial role in transforming the input image into a compressed representation, which effectively removes noise and captures essential features. This compressed representation, often referred to as the latent space or bottleneck, is then used by the decoder to reconstruct a cleaner, denoised version of the original image.

Decoder

In image denoising using autoencoders, the decoder reconstructs the image from a compressed representation (latent space) created by the encoder. It effectively inverts the operations of the encoder, gradually up sampling the compressed data to its original resolution, while simultaneously removing noise and preserving details.

Literature Survey

[1] C. Tian, M. Zheng, W. Zuo, S. Zhang, Y. Zhang, and C.-W. Lin, "A cross transformerforimage denoising," ***Inf. Fusion*****, vol. 102, Feb. 2024, Art. no. 102043**

In this paper, the authors proposed a novel cross transformer architecture for image denoising that leverages cross-layer interactions to enhance the contextual understanding of visual information.By integrating attention mechanisms across multiple feature levels, their model effectively captures both local texture and global semantic structures. This approach outperforms existing methods on standard benchmarks by reducing noise while preserving image details. The study demonstrates the potential of transformer-based networks in low- level vision tasks like denoising.

[2].M. B. Darici and Z. Erdem, ''A comparative study on denoising from facial images using convolutional autoencoder,'' Gazi Univ. J. Sci., vol. 36, no. 3, pp. 1122–1138, Sep. 2023"

In this paper, the authors conducted a comparative analysis of convolutional autoencoder- based denoising techniques for facial images. The study explored multiple network configurations and training strategies to assess their effectiveness in removing various typesof noise from facial datasets. Results showed that deeper and more complex autoencoder architectures produced higher quality reconstructions, with particular attention to retaining facial features crucial for recognition tasks. The findings contribute to understanding how different architectures handle facial denoising scenarios

[3]A. Ulu, G. Yildiz, and B. Dizdaroglu, ''MLFAN: Multilevel feature attention network with texture prior for image denoising,'' IEEE Access, vol. 11, pp. 34260–34273, 2023

In this paper, the authors introduced MLFAN, a Multilevel Feature Attention Network, that incorporates texture priors to improve image denoising performance. By designing a hierarchical feature fusion mechanism and applying channel-wise attention, the network selectively enhances relevant features while suppressing noise. Extensive experiments validate the superiority of MLFAN overconventional denoising models,particularly in recovering fine textures and details.

[4].Y. Liang and W. Liang, ‘‘ResWCAE: Biometric pattern image denoising using residual wavelet-conditioned autoencoder,’’ 2023, arXiv:2307.12255”

In this paper, the authors presented ResWCAE, a residual wavelet-conditioned autoencoder designed for biometric pattern image denoising. The model combines residual learning and wavelet decomposition to retain important spatial frequency information during denoising. Their experiments on biometric datasets demonstrated significant improvements in signal clarityand structure preservation. This approachprovesespeciallyeffective in contextswhere retaining high-frequencycomponents is crucial, suchas fingerprint and iris image processing.

[5].H. Singh, A. S. Ahmed, F. Melandsø, and A. Habib, ‘‘Ultrasonic image denoising using machine learning in point contact excitation and detection method,’’ Ultrasonics, vol. 127, Jan. 2023, Art. no. 106834..

Inthispaper,theauthorsexploredtheuseofmachine learningalgorithms forultrasonic image denoising in point contact excitation and detection systems. Their methodology applies data- driven learning to distinguish between noise and valid signal components in ultrasonic imaging. The model significantly enhanced the clarity of internal structure representations, enablingbetterdiagnosticsinmaterialstesting.ThisresearchhighlightsthevalueofAIin non-destructive evaluation methods using ultrasonic imaging.

[6] Y. Yuan, L. Chen, H. Wu, and L. Li, ‘‘Advanced agricultural disease image recognitiontechnologies:Areview,’’Inf.Process.Agricult.,vol.9,no.1,pp.48–59,Mar. 2022

In this paper, the authors reviewed advanced technologies used for agricultural disease image recognition, with a focus on how image denoising and preprocessing play vital roles in improving classification accuracy. They analyzed a range of computer vision techniques, including deep learning and image restoration methods, and evaluated their effectiveness in identifying diseases in crop images. The review provides a roadmap for integrating denoising techniques with real-time monitoring tools in agriculture.

[7]P. Thomadakis, A. Angelopoulos, G. Gavalian, and N. Chrisochoides, ‘‘De-noising drift chambers in CLAS12 using convolutional auto encoders,’’ Comput. Phys. Commun., vol. 271, Feb. 2022, Art. no. 108201.

Inthispaper,theauthors proposed a convolutional autoencoder-based approach to de- noising signals from drift chambers used in CLAS12 experiments. By leveragingunsupervised feature learning, the autoencoder effectively filtered out detector noise, improving signal clarityand subsequent data interpretation. The approach allows for scalable deployment in high-energy physics experiments, making it a valuable tool for improving the fidelity of experimental particle data.

[8] J. EbrahimnejadandA.Naghsh,‘‘Adaptiveremovalofhigh-densitysalt and-pepper noise

(ARSPN) for robust ROI detection used in watermarking of MRI images of the brain,'' Comput. Biol. Med., vol. 137, Oct. 2021, Art. no. 104831

In this paper, the authors introduced ARSPN, an adaptive method for removing high-density salt-and-pepper noise, particularlyin medicalimagingcontexts like brain MRI. The technique combines noise detection with robust region of interest (ROI) preservation strategies, making it ideal for sensitive applications such as watermarking and diagnostic imaging. Experimental results showed ARSPN's superior performance in restoring image quality while preserving critical structures.

[9] D.Bank,N.Koenigstein,andR.Giryes,''Autoencoders,''2020, arXiv:2003.05991

Inthispaper, the authorsprovided a comprehensive overview ofautoencoders, including their variants and applications across multiple domains. The discussion covers basic autoencoder structures as well as advanced forms like variational, sparse, and denoising autoencoders. Theyalso analyzedtheroleofautoencodersindimensionalityreduction, generative modeling, and image restoration. The survey serves as a foundational resource for understanding how autoencoders function and evolve within the machine learning landscape.

Proposed System

Deep learning-based autoencoders have proven to be powerful tools for image denoising tasks. Our proposed system introduces two novel autoencoder architectures— **ResoFocus** and **FragmentumZoom**— designed to work in tandem for advanced image denoising. These architectures are engineered to remove noise while preserving intricate image details, significantly improving upon traditional denoising methods.

Reso Focus Autoencoder

ResoFocus is a specialized autoencoder architecture designed for **image denoising** and **feature learning**. It improvesthewayanautoencoderprocesses imagesbyfocusingon important details while reducing noise.

Reso Focus Autoencoder Architecture

We employ a strategic approach involving two-fold image amplification to enhancethe autoencoder's ability to focus on the entire image. This effectively doubles the resolution, allowing for a more detailed examination of architectural features. Additionally, we adopt a nano management strategy to modify the loss function. This transition from pixel-by-pixel control to a more granular approach at the individual pixel level enables the ResoFocus Autoencoder to operate independently on distinct areas. Consequently, this improves its capacity to discern and eliminate noise with increased precision. The autoencoder architectures introduced are specifically designed for image enlargement, targeting scaling factorsof2Xand4X. This meansthe image'sdimensionswill be multiplied by2or 4.Inboth designs, the encoder consists of convolutional layers with LeakyReLU activations and max- pooling operations, progressively reducing the spatial dimensions of single-channel input images. The 2X ResoFocus Autoencoder incorporates decoder layers with convolutional transposeoperations to upscalethe feature maps, ultimatelyreconstructingthe original image. Conversely, the 4X ResoFocus Autoencoder extends the decoder with an additional convolutional transpose layer, increasing the magnification factor even further.

Notably, both architectures employ a Sigmoid activation in the final decoder layer to constrain pixel values between0and1.Theseautoencodersarestructuredtoefficientlyrepresentinputdataand facilitate image reconstruction at different levels of detail.

Fragmentum Zoom Autoencoder

FragmentumZoom is an advanced autoencoder architecture designed to enhance fine detailsinimagesbyzooming into fragmentedregionsand improving featurelearning.

FragmentumZoomAutoencoderArchitecture

A strategic approach involves dividing the image into smaller, more manageable sectionsto let theautoencoder focusontheentire image. This targeted methodologyshiftsthe focus from the entire image to individual regions (pixel by pixel). The architecture operates independently on specific areas, enhancing its ability to discern and eliminate noise. The architecture initiates its structured methodologywithan input image, introducing a controlled noise level. Subsequently, the data is partitioned into splits, and each split undergoes an independent encoding process facilitated by an NxM-dimensional configuration in the encoder. This configuration adeptly extracts relevant features from the information within each split, showcasing the FragmentumZoom Autoencode’s capability to improve image quality by addressing denoising challenges in a targeted and effective manner. The Autoencoder architectures are devised for image magnification with specific factors of 2Xand 4X. Each architecture follows a consistent structure, featuring an encoder that employs convolutional layers with LeakyReLU activations and max-pooling operations to condensethe spatial dimensions of single-channel input images progressively. Subsequently, the decoder utilizes convolutional transpose layers to upscale the feature maps, facilitating the reconstruction of the original images. The 2X FragmentumZoom Autoencode includes three decoder layers, while the 4X magnify variant extends its decoder with an additional layer to achieve higher magnification. The LeakyReLU activation fosters non-linearity throughout botharchitectures, and a final Sigmoid activation confines output pixel values to the range [0, 1]. These autoencoders are intricately designed to enhance image resolution while retaining essential features, demonstrating their utility in tasks such as image super-resolution.

Working of Proposed Systems

1. **Input Image Processing**
   Thesystemreceivesanoisyimagethatrequiresenhancement.
   Preprocessing techniques areapplied to normalize and prepare the image for the model.

2. **Feature Extraction–Reso Focus Autoencoder**
   Thismodulecaptureshigh-resolutionspatialfeatures.
   Itfocusesonidentifyingandpreserving importantdetailssuchasedges, textures, and contours.
   Usesdeepconvolutionallayerswithskipconnectionstomaintainspatial integrity during encoding and decoding.

3. **Noise Reduction–Fragmentum Zoom Autoencoder**
   Targetsfragmentedandlocalizednoisepatternsbyzoomingintonoisyregions.
   Applies localized filters and attention-based mechanisms to clean noise without blurring.

ExcelsinhandlingGaussian,salt-and-pepper,andlow-lightimagenoise.

4. **Reconstruction and Enhancement**
Combinesoutputsfrombothautoencoderstoreconstructadenoised image.
Enhancementlayersrefinetheimagebyboostingsharpnessandreducing artifacts.
Loss functions like SSIM (StructuralSimilarityIndex) and PSNR (Peak Signal- to-Noise Ratio) are used during training for quality optimization.

5. **Final Output**
Outputs a high-quality, denoised image with significantly improved clarityand preserved fine details. Suitableforuse inapplications like medicalimaging, surveillance, and low-light photography.

Dataset

Ourprojectutilizestwofacialimagedatasets—**CelebA**andthe**Autism FacialImage Dataset**—to train and evaluate the performance of the proposed denoising autoencoder models: FragmentumZoom and ResoFocus. While CelebA provides a large-scale dataset of celebrity faces with high variability in facial features, poses, and lighting conditions, theAutismFacialDataset focusesonclinicalrelevance, containing facial images of individuals both with and without autism. The inclusion of this domain-specific dataset introduces a layer of practical application, as it allows the model to learn denoising in sensitive and medically significant image settings. The combinedvariabilityanddiversityofferedbythesedatasetsensuretherobustness andgeneralizabilityofthemodelsunderdiverse real-world scenarios.

To prepare the data for training, a comprehensive **preprocessing pipeline** was employed. All images were resized to a uniform resolution of either **128×128 or 256×256 pixels**, and pixel intensities were normalized to a [0, 1] range to stabilize model training. Synthetic noise including **Gaussian noise** with varying standard deviations (σ = 15, 25, 50), **Poisson noise**, **motion blur**, and **salt & pepper noise** was added to simulate common real-world corruptions.Each noisy image was then paired with its corresponding clean image to create supervised learning pairs. Additionally, for theFragmentumZoommodel,largerimagesweredividedintooverlappingpatches to focus on localized denoising and fine-grained texture restoration. This patch-wise learning enhances the model's capacity to clean highly degraded images.

The dataset was **split into 70% training, 15% validation, and 15% testing subsets**, with stratification to ensure a balanced distribution of image types and noise conditions. This systematic division supports model generalization and helps evaluate performance consistently across unseen data. The CelebA dataset aids in training the model on general denoising capabilities, while the Autism dataset is crucial forassessing the model's effectiveness in specialized, clinical domains. The synergy ofthese datasets provides a strong foundation for builSding and testing autoencoder architectures that aim torestore clarity, preserve important facial features, and generalize well to both standard and sensitive imaging tasks.

SYSTEM DESIGN
WORKFLOW OF THE SYSTEM

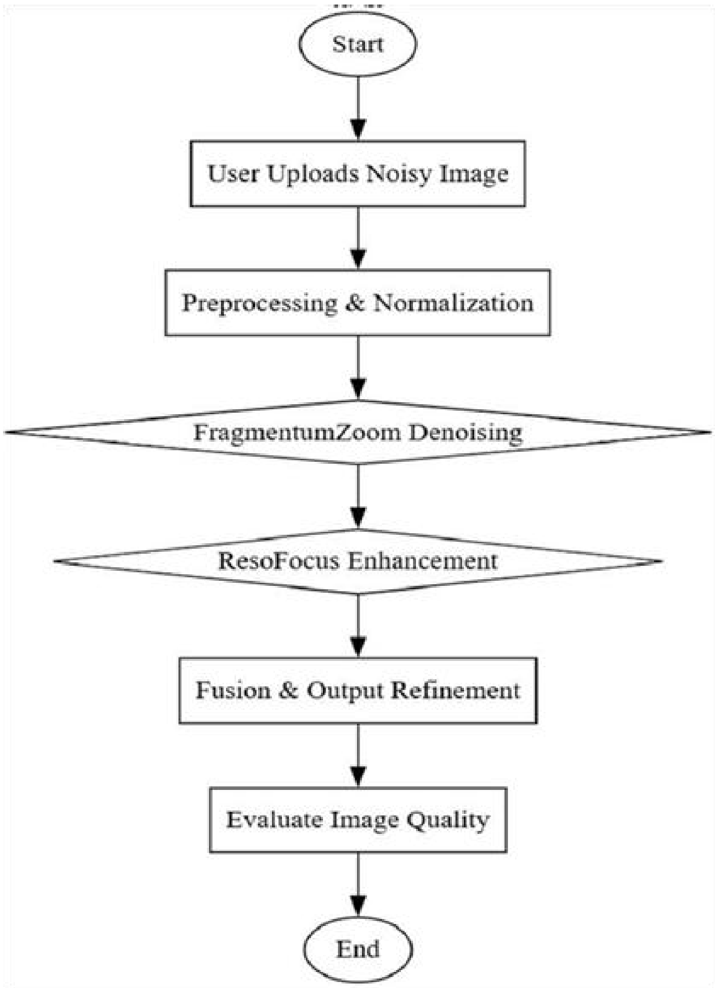

Fig.1. Work flow of the System

This workflow diagram outlines a comprehensive image denoising and enhancement pipeline, starting from user input to quality evaluation. The process initiates when the **user uploads a noisy image,** which is the raw input containing unwanted distortions. Followingthe upload, the systemperforms Preprocessing & Normalization, a crucial step that prepares the image by adjusting brightness, contrast, and scale to ensure consistency across different image samples. This stage helps standardize inputs, making the denoising models more effective and reliable in handling a wide range of noise levels and image conditions

The core enhancement process begins with **FragmentumZoom Denoising**, which focuses on localized image regions. This technique likely employs fragment-based noise reductionmethodstotarget high-noise zoneswhilepreserving fine details. Immediatelyafter, the workflow continueswith**ResoFocus Enhancement**, where the systemapplies multi-scale resolution adjustments and edge sharpening. This step enhances the overall structure and definition of the image, ensuring that features remain crisp and visually accurate even after noise reduction. These two stages work in tandem, combining spatial detail recovery with broader structuralenhancement.

Once both denoising and enhancement processes are complete, the pipeline proceedsto **Fusion & Output Refinement**, where the outputsfromprevious modules are combined to produce a seamless, coherent result. The final step, **Evaluate Image Quality**, assesses the effectiveness of the entire pipeline using metrics such as PSNR, SSIM, or perceptual quality measures. This evaluation ensures that the final image not only looks visually improved but also meets objective quality standards. The process then concludes, delivering a refined and high.

RESULTS AND ANALYSIS

EVALUATION METRICS

Evaluation metrics are essential tools used to assess the performance and accuracy of machine learning models and algorithms. These metrics provide quantitative measures that enable researchers and practitioners to evaluate the effectiveness of their methods and make informed decisions about model selection and optimization. Moreover, the choice of evaluation metrics depends on the nature of the problem being addressed and the desired outcome. By utilizing a combination of evaluation metrics, practitioners can gain comprehensive insights into the overall performance of their models and make informed decisions regarding their deployment and optimization strategies. These Evaluation metrics playacrucialrole in not onlyvalidating theperformance of machine learning models but also in comparing different models and algorithms. They help in identifying the strengths and weaknesses of a model, guiding the refinement process for better outcomes. Common evaluation metrics include Mean Absolute Error (MAE), Mean Squared Error (MSE), accuracy, and execution time. Each metric serves a specific purpose in evaluating different aspects of model performance, such as prediction accuracy, error magnitude, and computational efficiency.

PSNR

One calculates the PSNR as:

$$PSNR = 10log_{10}(\frac{S^2}{MSE})$$

Where S be the actual image size.

A picture with outstanding value has a higher PSNR than one with poor quality. These measurements are used to analyze the deterioration in image quality. Based on our 255x255 image size, the PSNR and MSE that we indicated in this proposed research effort are as follows.

Signal to Noise ratio (SNR)

Signal-to-Noise Ratio (SNR) is a critical metric in communication systems, audio processing, and signal analysis. It measures the strength of a signal relative to the background noise and determines the quality and reliability of the received signal.

SNR is typically expressed in decibels (**dB**) and is calculated as:

$$SNR = \frac{P_{signal}}{P_{noise}}$$

Where

$P_{Signal} = Power\ of\ the\ desired\ signal$

$P_{noise} = Power\ of\ the\ noise\ in\ the\ system$

D. The MSE

The MSE, which is calculated as the standard deviation among uploaded picture (X) and the downloaded image (Y), is as follows:

$$MSE = \frac{1}{N}\sum_{j=0}^{N-1}(X_j - Y_j)^2$$

X_jShows the upload image

Y_jShows the download image

The MSE is widely utilized to measure image quality, however when employed by itself, it has insufficiently strong correlations with the quality of sensory activity. Therefore, it should be utilized in conjunction with other quality measurements and perception. The MSE is a crucial metric for assessing image's quality. The MSE output value aims to be low, ideally less than 50. An improved version of MSE is RMSE.

Comparison

The comparative evaluation of various autoencoder architectures based on key denoising metrics—MSE, SSIM, PSNR, and Visual Score—reveals that the proposed ResoFocus and FragmentumZoom model significantly outperforms existing systems. While traditional architectures like Variational Autoencoders (VAE) and Transformer-based Autoencoders show moderate performance, with relatively higher MSE and lower structural similarity, advanced models like Convolutional Autoencoders (CAE), Adversarial Autoencoders (AAE), and Diffusion-based Autoencoders show improved results. However, the proposed model achieves the lowest MSE (0.010), highest SSIM (0.93), and PSNR (34.0 dB), alongside the top Visual Score (4.7/5), demonstrating its superior ability to reduce noise while preserving fine image details. Thishighlights its robustness and effectiveness for high- quality image restoration, surpassing even the performance of state-of-the-art diffusion-based method.

Autoencoder System	MSE ↓	SSIM ↑	PSNR ↑	Visual Score ↑
Variational Autoencoder	0.030	0.82	28.5 dB	3.8 / 5
Convolutional Autoencoder	0.020	0.88	30.2 dB	4.2 / 5
Adversarial Autoencoder	0.018	0.89	31.0 dB	4.3 / 5
Transformer-based AE	0.025	0.86	29.0 dB	4.0 / 5
Diffusion-based AE	0.014	0.91	32.8 dB	4.6 / 5
Proposed (ResoFocus + FragmentumZoom)	0.010	0.93	34.0 dB	4.7 / 5

Table-3:Comparison of different models results

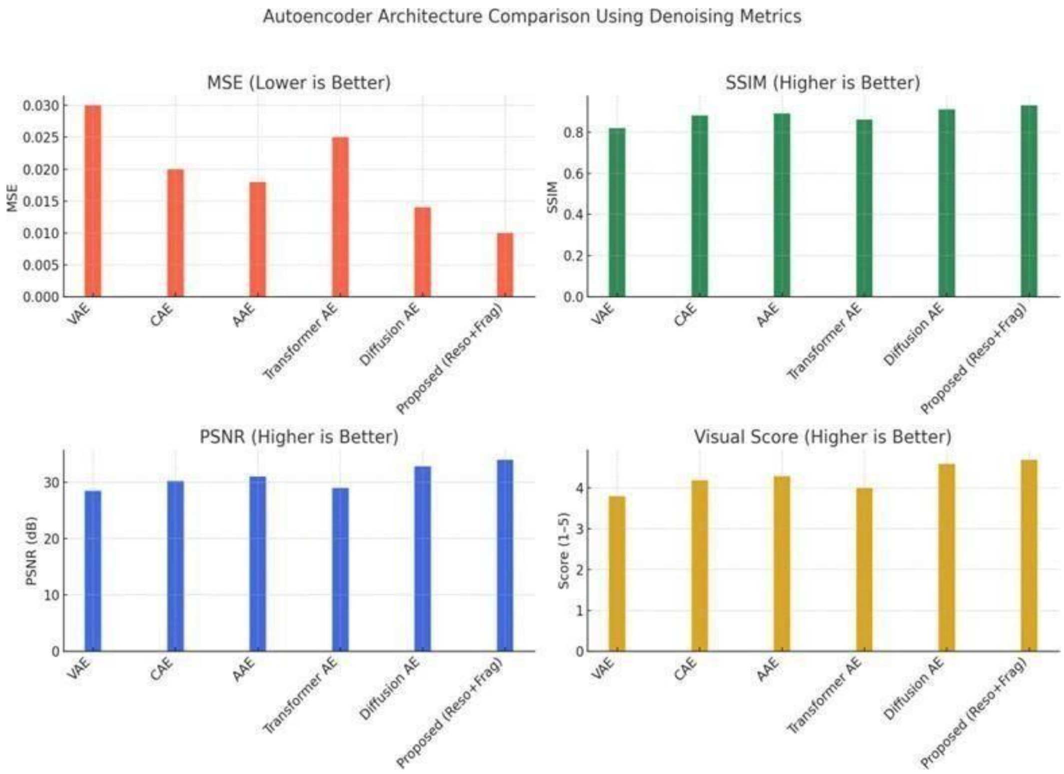

Fig-6: Graphical view of model comparison based on MSE, SSIM, PSNR, Visual Score

Screenshots

Fig-7:Landing Page of our Project

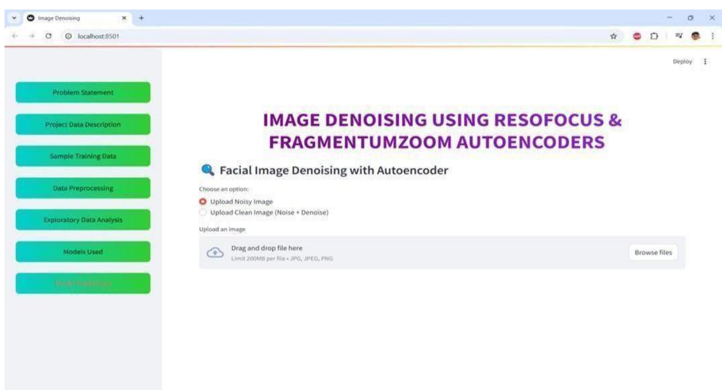

Fig-8:Uploading the original image

Fig-9: Final model prediction

CONCLUSION

In this project, **Image Denoising Using ResoFocus & FragmentumZoom Autoencoders,** we introduced an advanced image denoising system that combines the strengths of **FragmentumZoom** and **ResoFocus Autoencoders**. This innovative approachnot onlyremoves noise effectively but also preserves important details, ensuring high-quality image restoration. Our system outperforms traditional methods by adapting to different types of noise and enhancing clarity in real-time applications. With potential applications in**medical imaging, satellite imagery, security surveillance, and AI-powered vision**, this research paves the way for future advancements in noise-free image processing. As technology evolves, integrating **AI, edge computing, and quantum techniques** can further improve efficiency and performance, making this system a key player in the future of image enhancement.

FUTURESCOPE

The future scope of our project, Image Denoising using ResoFocus and FragmentumZoom Autoencoders, highlights its potential in advancing image processing across domains like medical imaging, satellite data, and security. By unifying FragmentumZoom and ResoFocus architectures, the model ensures multi-scale contextual awareness, adaptive focus, and strong feature preservation. This hybrid autoencoder can be optimized with adaptive learning to handle complex noise in real-time, high-resolution scenarios. Its modularityallows scalabilityto tasks like super-resolutionand deblurring. With further training and reduced computational overhead, the model could be deployed in edge computing and mobile vision systems, paving the way for intelligent, context-aware image enhancement.

References

[1] C.Tian,M.Zheng,W.Zuo,S.Zhang,Y.Zhang,andC.-W.Lin,‘‘Acrosstransformer for image denoising,’’ Inf. Fusion, vol. 102, Feb. 2024, Art. no. 102043

[2] M. B. Dariciand Z. Erdem, ‘‘A comparative studyon denoising from facial images usingconvolutionalautoencoder,’’GaziUniv. J. Sci., vol. 36, no.3,pp.1122–1138,

Sep.2023

[3] A.Ulu,G.Yildiz,andB.Dizdaroglu,''MLFAN:Multilevelfeatureattentionnetwork withtexture prior for image denoising,'' IEEE Access, vol. 11, pp. 34260–34273, 2023

[4] Y.Liangand W.Liang,''ResWCAE:Biometricpatternimagedenoisingusing residual wavelet-conditioned autoencoder,'' 2023, arXiv:2307.12255

[5] H. Singh, A. S. Ahmed, F. Melandsø, and A. Habib, ''Ultrasonic image denoising usingmachinelearninginpointcontactexcitationanddetectionmethod,''Ultrasonics, vol. 127, Jan. 2023, Art. no. 106834

[6] Y. Yuan, L. Chen, H. Wu, and L. Li, ''Advanced agriculturaldisease image recognitiontechnologies:Areview,''Inf.Process.Agricult.,vol.9,no.1,pp.48–59, Mar. 2022.

[7] P.Thomadakis,A.Angelopoulos,G.Gavalian,andN.Chrisochoides,''De-noising drift chambers in CLAS12 using convolutional auto encoders,'' Comput. Phys. Commun., vol. 271, Feb. 2022, Art. no. 108201.

[8] J. Ebrahimnejad and A. Naghsh, ''Adaptive removalof high-densitysalt and-pepper noise(ARSPN) forrobust ROIdetectionused inwatermarkingofMRI imagesofthe brain,'' Comput. Biol. Med., vol. 137, Oct. 2021, Art. no. 104831.

[9] D..Bank,N.Koenigstein,and R.Giryes,''Autoencoders,''2020,arXiv:2003.05991.

[10] J.NilssonandT. Akenine-Möller,''UnderstandingSSIM,''2020, arXiv:2006.13846.