Comparison of Deep Learning Algorithms with Different Activation Functions for Brightness Image Enhancement

M. Fatih Yilmaz *fthyilmaz53@gmail.com*
Sheriff Latif LLC.
Fremont,CA, USA

Bekir Karlik *bkarlik@epoka.edua.al*
Computer Engineering
Epoka University Albania
Tirana-1000, Albania

Abstract

The training and refinement of deep learning-based image enhancement models are vital to their advancement. Researchers have examined various optimization methods, including adversarial training, multi-task learning, and perceptual loss functions, to enhance the quality and consistency of the improved results. This research explores brightness image enhancement techniques utilizing Autoencoder and Convolutional Neural Networks (CNN). Furthermore, the most appropriate activation functions for both approaches have been analyzed. Within this framework, functions like sigmoid, tanh, Softmax, ReLU, and leaky ReLU have been evaluated. The results suggest that tanh and leaky-ReLU are the most efficient activation functions for autoencoder neural networks, whereas sigmoid and SoftMax are the most appropriate for CNNs in image enhancement applications.

Keywords: Deep Learning, Image Enhancement, Activation Function, Loss.

1. INTRODUCTION

Improving images is an essential endeavor in the domains of computer vision and digital image processing, concentrating on boosting the visual appeal and clarity of digital pictures. Traditional methods, including histogram equalization, contrast adjustment, and edge enhancement, have been widely applied; however, they frequently encounter difficulties when managing complex distortions and artifacts(Mustafa et al., 2019).

For the past thirty years, AI-driven image compression and enhancement have been explored. At first, Karlik et al. (1996) presented a parallel autoencoder framework for image compression, contrasting it with conventional compression methods. As illustrated in Figure 1, this model features two parallel Multi-Layered Perceptron (MLP) structures. The first MLP is responsible for coding with a target (or desired) output, where the input and the desired output consist of similar data. The second MLP focuses on encoding (or reconstruction) and is validated and refined by attempting to recreate the original input from the encoding. Both MLPs consist of layers that are completely interconnected. In this framework, the system utilizes the output from the midpoint of the weights, which indicates the compression ratio (Karlik, 2000).

Subsequently, Karlik enhanced an alternative model known as Hierarchical Finite State Vector Quantization (HFSVQ) along with a neural network (Karlik, 2006). This algorithm achieved the nonlinear restoration of diffraction-limited medical images while simultaneously performing quantization. This approach can be applied to both 2D image compression and 1D biomedical signal compression (Kocyigit, 1999).

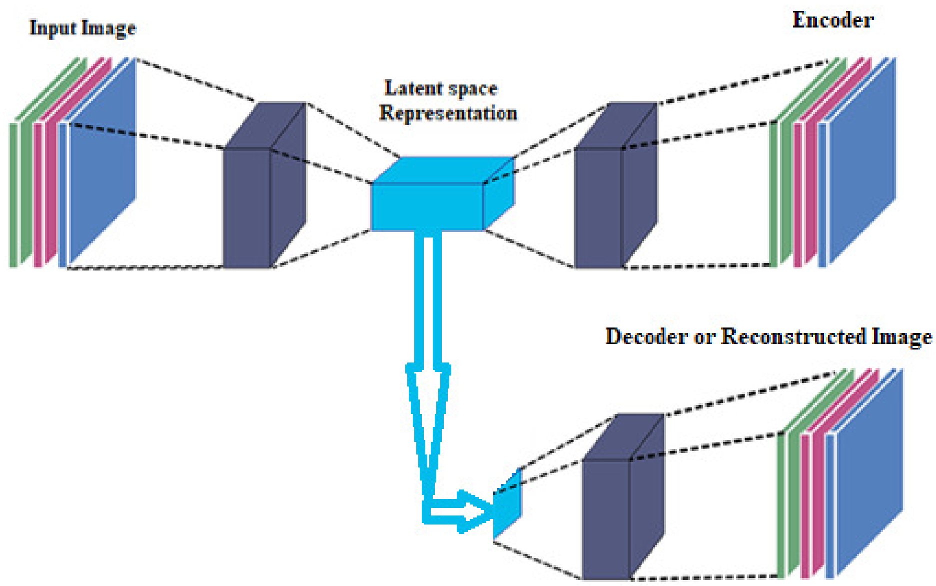

FIGURE1: The architecture of parallel autoencoder neural networks.

As the domain of image enhancement progresses, there is an increasing focus on creating models that are customized for specific areas, such as document images. Upcoming research might investigate the incorporation of domain-specific insights, sophisticated data augmentation strategies, and innovative network designs to further enhance the efficiency of image enhancement systems (Khan, 2023; Tensmeyer & Martinez, 2017; Kirsten, 2021; and DeVries & Taylor, 2017).

Comprehensive evaluations have contrasted the effectiveness of deep learning-driven image enhancement models with traditional computer vision methods. The findings typically illustrate the advantages of deep learning techniques, as the top-performing architectures yield visually stunning and high-quality enhanced images (Kirsten, 2021 and Gupta, 2020). The emergence of deep learning has transformed the realm of image enhancement, with convolutional neural networks and autoencoder models showcasing exceptional results across various enhancement tasks (DeVries & Taylor, 2017 and Long, 2014). Particularly in the field of medical imaging, transfer learning has demonstrated itself to be a valuable method for utilizing models that have already been trained(Thomas, 2021). The significant domain disparity between natural images and medical images often obstructs the direct use of models trained on natural image datasets. These deep learning-based techniques are capable of learning intricate hierarchical feature representations, allowing them to effectively tackle the challenges faced by traditional methods (DeVries & Taylor, 2017).

Recent developments have also investigated the fusion of convolutional and autoencoder frameworks for improving image quality. These combined models utilize the advantages of both methodologies, where the convolutional layers effectively capture intricate visual characteristics and the autoencoder framework facilitates proficient representation learning and reconstruction. Such integrated models have yielded encouraging outcomes in applications such as super-resolution, artifact elimination, and improved detail retention (Gupta et al., 2020). As discipline progresses, the collaborative integration of these deep learning strategies is anticipated to lead to further advancements in image enhancement abilities. Convolutional neural networks are especially well-equipped for image enhancement endeavors due to their capacity to extract and amalgamate complex visual attributes. The convolutional layers within these networks are capable of identifying low-level characteristics, like edges and textures, alongside higher-level semantic elements, which can be utilized to enhance the quality of the output (DeVries & Taylor, 2017). Studies indicate that the depth and intricacy of CNN architecture are pivotal to its efficacy in image enhancement. More profound networks with additional convolutional layers can capture increasingly abstract and task-specific attributes, resulting in superior enhancement results. However, as the intricacy of the network increases, the likelihood of overfitting rises, making it essential to apply appropriate regularization techniques. Furthermore, deep learning-based techniques for image enhancement can also be employed to improve the brightness of low-brightness remote sensing imagery (Hu, et al., 2021).

Based on the literature reviewed, convolutional neural networks and models utilizing autoencoders have demonstrated encouraging outcomes in enhancing image quality, exceeding the capabilities of conventional image processing methods (Bhardwaj, 2017;Anwar & Li,2020). For example, research conducted by Bhardwaj et al. in 2017 revealed that techniques based on deep learning, like generative adversarial networks, were proficient in producing visually striking and realistic enhanced images, outperforming traditional super-resolution techniques (Chakraborty et al., 2024]. Additionally, another study assessed multiple cutting-edge deep neural network frameworks for enhancing document images and discovered that the top-performing models delivered significant improvements compared to established traditional computer vision approaches. These results indicate that image enhancement methods rooted in deep learning can effectively tackle a range of image degradation challenges, including noise, blurriness, and low resolution while offering a more automated and adaptable solution than traditional strategies.

2. MATERIALS ANDMETHODS

Adversarial training can be employed to produce more realistic and visually stunning, improved images by integrating a discriminator network that learns to tell apart genuine samples from those that are artificially generated. Conversely, multi-task learning enables the model to learn several tasks simultaneously, leading to better generalization and improved skills.

Assessing the effectiveness of image enhancement models is a challenging endeavor, as it encompasses both objective and subjective criteria. Objective metrics like PSNR, SSIM, and LPIPS can deliver quantitative evaluations of enhancement quality, while human perceptual assessments can provide crucial insights into the visual attractiveness and practical utility of the enhanced results. Researchers have also investigated the creation of domain-specific evaluation metrics that are more closely aligned with the unique needs of particular applications, such as medical imaging or satellite imagery. Moreover, efficient preprocessing techniques like normalization, resizing, and contrast adjustments can enhance the input data and support the learning process of neural networks.

The quality and variety of the training data are vital factors that significantly impact the performance of deep learning-driven image enhancement models. To address this challenge, researchers have explored a range of data augmentation techniques, including random cropping, flipping, and the addition of noise, to expand the pool of training data and improve the models' ability to generalize.

Additionally, integrating image enhancement techniques with other computer vision activities, like object detection or semantic segmentation, can lead to more holistic and collaborative approaches to image analysis and understanding. By focusing on these prospective paths, the field of deep learning-driven image enhancement can continue to make significant advancements, providing effective and applicable solutions across a wide range of uses.

As the adoption of deep learning-driven image enhancement grows, it becomes crucial to reflect on the ethical ramifications of these technologies. Possible issues include the risk of biased or unfair outputs, the danger of misleading manipulation of visual data, and the potential for the improper use of enhanced images in various contexts. Researchers and professionals must tackle these ethical challenges by establishing robust guidelines, creating ethical frameworks, and formulating responsible implementation strategies to guarantee the safe and reliable application of image enhancement models.

The improved images produced by models utilizing deep learning techniques have numerous applications, such as:

- Medical imaging: Enhancing the quality of medical visuals can support diagnosing and planning treatments (Gupta et al., 2020).
- Document digitization: Improved document images can lead to enhanced optical character recognition and easier retrieval.
- Surveillance and security: Enhanced monitoring footage can offer better identification of objects and people.
- Historic preservation: Image enhancement can assist in restoring and safeguarding important historical documents and photos.

These applications demonstrate the extensive influence and practical advantages of advancements in deep learning-driven image enhancement (DeVries & Taylor, 2017; Gupta et al., 2020; Coghlan & Quinn, 2023; Hernández-García & König, 2018; Taylor & Nitschke, 2017; and Khosla & Saini, 2020).

This research has employed two specific deep-learning models for image enhancement and has analyzed their outcomes.

2.1 Autoencoder Neural Networks for Image Enhancement

Autoencoder neural networks, in contrast, are defined by their encoder-decoder framework, which facilitates effective feature representation and reconstruction. The encoder section of an autoencoder captures a condensed, latent representation of the input image, which is subsequently transmitted to the decoder to recreate the improved output. This approach allows autoencoders to create effective feature representations that highlight the most significant aspects of the input, making them especially useful for applications like noise suppression, artifact removal, and enhancing image resolution (Peng, et al., 2023). Generally regarded as an image compression technique, autoencoders are a form of multilayer neural network. Furthermore, autoencoder neural networks have demonstrated significant potential in the realm of image enhancement. These models consist of an encoder part that captures a concise representation of the input image and a decoder part that reconstructs the enhanced output. The encoder-decoder framework empowers autoencoders to learn efficient feature representations that can be utilized for various enhancement applications.

Recent studies have investigated the integration of generative adversarial networks with autoencoders for the enhancement of historical text images. These models are capable of effectively tackling issues caused by severe degradations, such as noise, fading, and artifacts, commonly encountered in old or damaged historical documents.

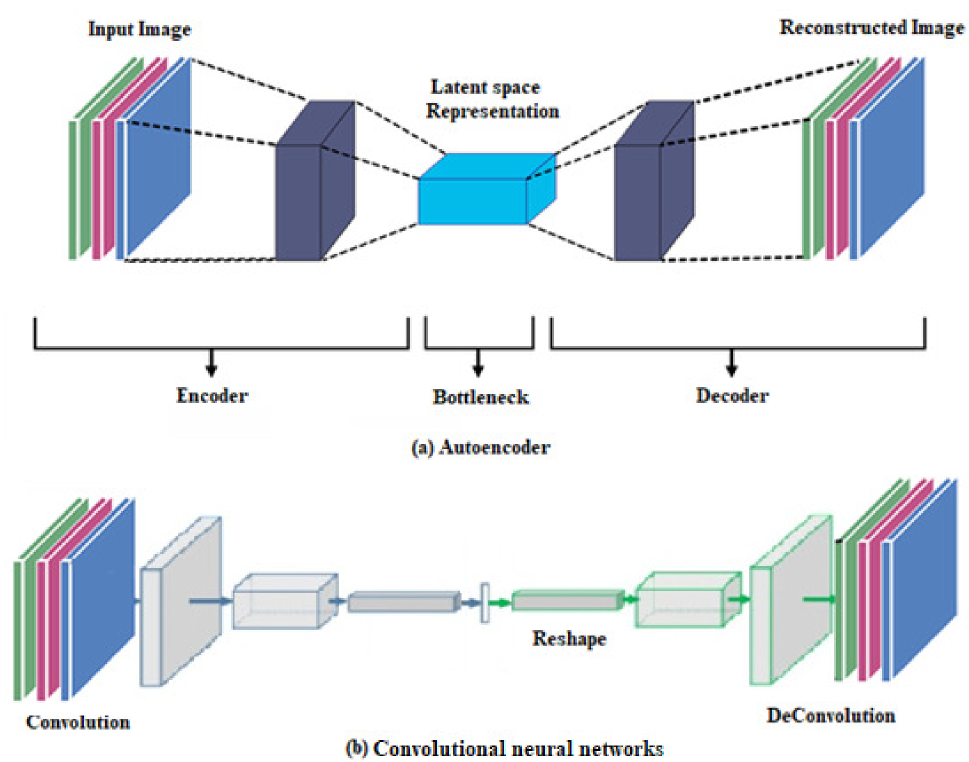

FIGURE 2: The architecture of autoencoder and conventional neural networks

Autoencoders are categorized into Denoising (or traditional) Autoencoder, Sparse Autoencoder, Deep Autoencoder, Contractive Autoencoder, Incomplete Autoencoder, Convolutional Autoencoder, and Parallel Autoencoder based on their respective applications. In this study, we utilized two different varieties of autoencoders. The Denoising Autoencoder consists of three key components: the encoder, the decoder, and the representation of the latent space. This network is capable of reconstructing the original image after it processes the input image and compresses it into its deepest hidden layer. As demonstrated in Figure 2 (a), it consists of multiple linear layers and is categorized as an unsupervised method because it functions using unlabeled data. The encoder identifies the essential features, the latent space serves as a storage for these features, and the decoder performs the reverse function of the encoder, reconstructing an image that closely resembles the original by employing the recognized features. Autoencoders find applications in signal and image processing for tasks such as denoising or data reconstruction.

A component used in the final stage of a neural network is referred to as a dense layer. Every node within the dense layer is intricately linked to the nodes of the following layer, from which it obtains the produced outputs. The nodes of the dense layer execute matrix-vector multiplication. For this operation to take place, the number of rows in the output vector from the previous layer must match the number of columns in the dense layer vector.

DECODER	ENCODER	LOSS
SIGMOID	SIGMOID	0.1194
SIGMOID	TANH	0.1233
SIGMOID	RELU	0.1729
SIGMOID	LEAKY-RELU	0.1022

SIGMOID	SOFTMAX	0.1495
RELU	SIGMOID	0.0118
RELU	TANH	0.0012
RELU	RELU	0.0128
RELU	LEAKY-RELU	0.0007
RELU	SOFTMAX	0.1496
LEAKY-RELU	SIGMOID	0.0109
LEAKY-RELU	TANH	0.0024
LEAKY-RELU	RELU	0.0004
LEAKY-RELU	LEAKY-RELU	0.0006
LEAKY-RELU	SOFTMAX	0.1495
TANH	SIGMOID	0.0091
TANH	TANH	0.0059
TANH	**RELU**	**0.000076**
TANH	**LEAKY-RELU**	**0.00039**
TANH	SOFTMAX	0.1495
SOFTMAX	SIGMOID	0.1217
SOFTMAX	TANH	0.1216
SOFTMAX	RELU	0.2851
SOFTMAX	LEAKY-RELU	0.2834
SOFTMAX	SOFTMAX	0.1496

TABLE 1: Performance results according to used activation function for 100 epochs.

In this research, we investigated various pairings of activation functions used in both the encoder and decoder layers, including Sigmoid, Rectified Linear Unit (RELU), Leaky-RELU, Hyperbolic tangent (TANH), and SoftMax. As shown in Table 1, the top results were achieved first with the TANH & Leaky ReLU combination, followed by the TANH & RELU pair, and lastly the Leaky RELU & RELU combinations. In contrast, the widely used SoftMax and Sigmoid functions exhibited lower performance compared to the other activation functions tested.

Figure 3 illustrates the training loss and validation loss for Autoencoder neural networks that employ Tanh and ReLu as the ideal combination. Figure 4 illustrates the training and validation losses for Tanh and Leaky_ReLu as the second-most effective combination.

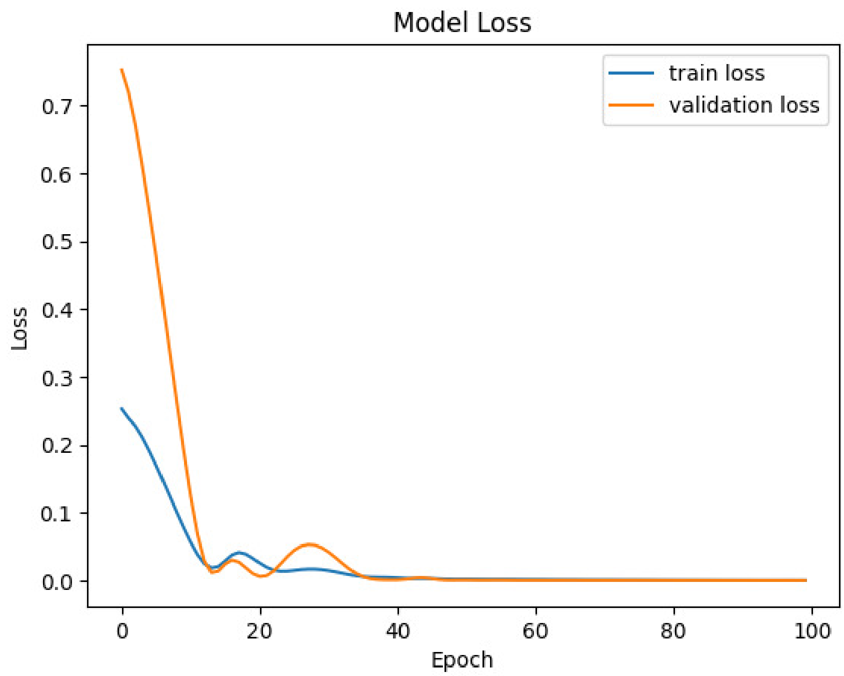

FIGURE 3: The loss graph according to the AE with tanh and leaky_ReLU.

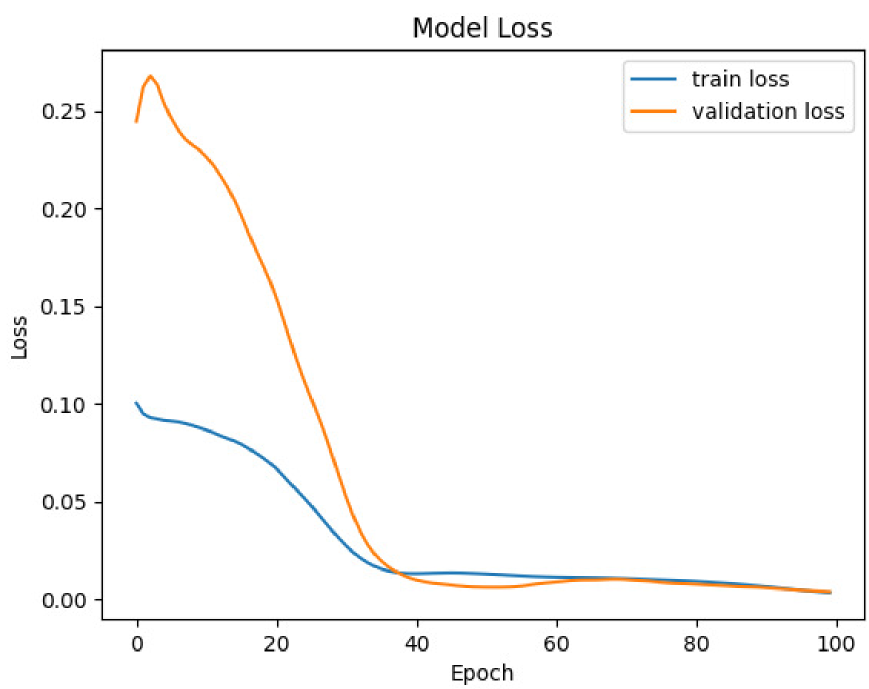

FIGURE 4: The loss graph according to the AE with Tanh and ReLU.

2.2 Convolutional Neural Networks for Image Enhancement

Convolutional neural networks (CNNs) have become an effective solution for enhancing images. These models are capable of learning to identify and integrate complex features from input images, allowing them to execute tasks like denoising, super-resolution, and artifact removal with remarkable precision. The capacity of CNNs to independently learn optimal features, instead of depending on manually adjusted parameters, has been a significant contributor to their success.

Numerous studies have assessed the efficacy of advanced CNN architectures for improving document images, showcasing their advantages over conventional computer vision methods (Kirsten et al., 2021). These models can adeptly manage a range of degradations, such as noise, fading, and inconsistent lighting, resulting in visually appealing and high-quality enhanced images.

Central to a Convolutional Neural Network are the convolutional layers, which draw inspiration from the biological visual system (O'Mahony et al., 2020). These layers operate as feature detectors, where each following layer becomes adept at recognizing more complex patterns within the input data. The convolutional layers consist of filters or kernels that slide over the input image, executing a dot product operation and generating a feature map. As the input information flows through the network, the feature maps progress to become increasingly abstract. The early layers detect fundamental elements like edges and shapes, whereas the later layers recognize more intricate patterns and objects (See Figure 2 (b)). Another crucial component of Convolutional Neural Networks is the pooling layers, which carry out a down-sampling operation on the feature maps.This process reduces the spatial size of the feature maps, enhancing the computational efficiency of the network and minimizing the likelihood of overfitting. The training procedure for a Convolutional Neural Network typically occurs via supervised learning, in which the network is given labeled data for training. The system then modifies the weights of its neurons based on the provided data, striving to reduce the gap between the predicted outcomes and the true labels. The success of Convolutional Neural Networks in visual applications is mainly due to their ability to learn and identify important features from the input data.

Table 2 shows the loss and validation loss results of CNN across the test images when sigmoid, RELU, LEAKY-RELU, SOFTMAX, and TANH are used as the dense activation function, respectively. As seen in Table 2, the lowest loss was found first for Sigmoid, and second for SoftMax. In other activation functions, loss was found to be very high. Interestingly, the higher-than-expected loss was observed in the ReLU activation function, which is the most used in the CNN method.

ACTIV FUNC.	SIGMOID	SOFTMAX	LEAKY-RELU	TANH	RELU
LOSS	0.0056	0.5747	3..2236	3.2237	3.2236

TABLE 2: Performance results according to used activation functions for 100 epochs.

Figures 5 and 6 illustrate the loss graph of the Convolutional Neural Network (CNN) over the test images, employing the Sigmoid and SoftMax activation functions, respectively. Consequently, it was noted that the enhancement outcomes on AE bright images using the combination of Tanh and ReLU outperformed the results from CNN. Nevertheless, the execution time for the CNN (39 seconds) is approximately four times quicker than that of the AE (1 minute and 53 seconds).

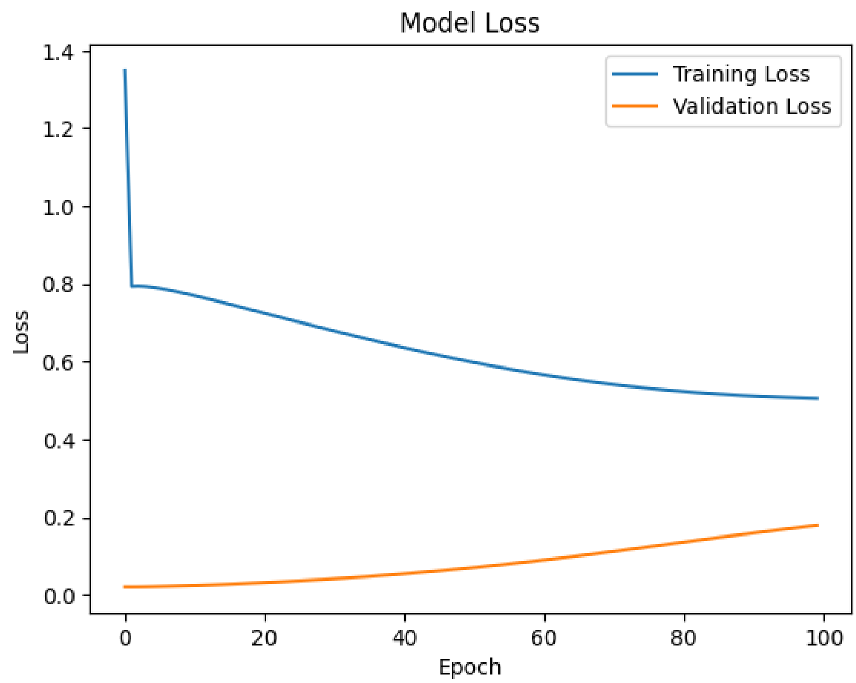

FIGURE 5: The loss graph according to CNN with Sigmoid.

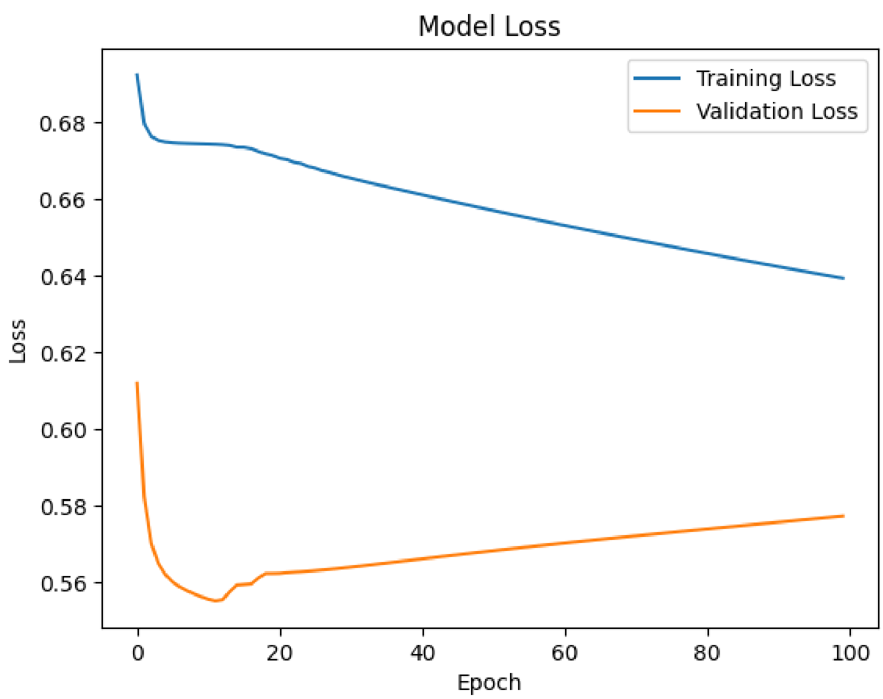

FIGURE 6: The loss graph according to CNN with SoftMax.

In this comparative analysis, the choice of suitable techniques for enhancing brightness in images through Autoencoder (AE) and Convolutional Neural Networks (CNN) was explored. Within this framework, commonly utilized activation functions such as sigmoid, tanh, SoftMax, ReLU, and leaky-ReLU were examined. The findings indicated that tanh, ReLU, and leaky-ReLU emerged as the most effective activation functions for autoencoder neural networks aimed at enhancing images. Conversely, Sigmoid and SoftMax emerged as the best activation functions for image enhancement when CNN was employed. Figure 7 illustrates the results of the image enhancement tests conducted on the selected brightness images using the b) Autoencoder (AE) with the combination of tanh and ReLU, as well as c) CNN utilizing the Sigmoid activation function.

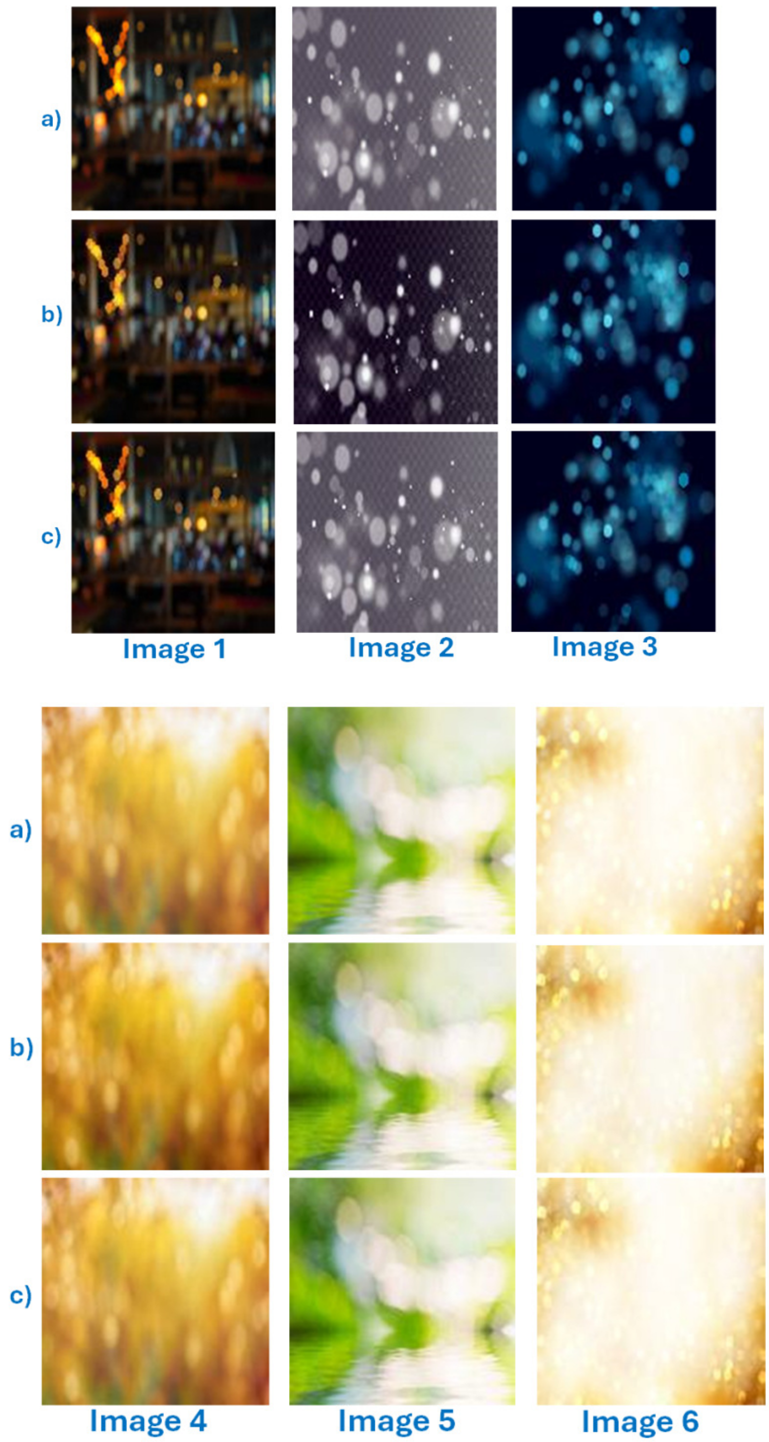

FIGURE 7: Image enhancement test results a) input images b) output images using the AE with tanh and ReLU combination, andc) CNN with Sigmoid.

3. CONCLUSION AND DISCUSSION

To summarize, the adoption of convolutional neural networks, parallel autoencoders, and the autoencoder framework has revolutionized the field of image enhancement. These sophisticated learning-driven techniques have been shown to surpass traditional computer vision approaches, enabling the effective handling of complex image degradations and the production of higher-quality enhanced outputs. As the field evolves, the incorporation of cutting-edge deep learning strategies, specialized knowledge in the field, and efficient approaches for data preprocessing and augmentation is expected to propel additional improvements in image enhancement(Kirsten, 2021; DeVries & Taylor, 2017,Gupta, 2020; Taylor & Nitschke,2017; and Tian, 2019). Our results align with the idea that while convolutional neural networks have been widely utilized for various image-processing applications, autoencoders have shown promise in areas such as image enhancement and noise reduction. The use of autoencoders, as evidenced by (Dosovitskiy & Brox, 2016; Augustauskas & Lipnickas, 2020; and Mushtaq, 2022) has led to improved image quality and enhanced performance in specific cases.

In this research, it was noted that employing distinct transfer functions in both the encoder and decoder parts of the autoencoder model proved advantageous. Within this framework, various combinations were explored, leading to the conclusion that the tanh function in the encoder paired with the ReLU function in the decoder yielded the most favorable outcome. In a related context, following trials with various activation functions for convolutional neural networks, it was determined that using sigmoid or softmax yielded better results. Nevertheless, the runtime for CNNs is roughly four times quicker than that of autoencoders based on the optimal models derived from both approaches. Future investigations could delve into the possible advantages of using autoencoders for restoration-focused tasks involving the enhancement of aged or fragmented images.

4. ACKNOWLEDGMENTS

The images utilized in this study were obtained from www.istock.com, with the following stock IDs: 539199982, 1423596003, 1494832020, 621116812, 1979495619, and 2160068696.

5. REFERENCES

Anwar, S., & Li, C. (2020, August 20). Diving deeper into underwater image enhancement: A survey. Elsevier BV, 89, 115978-115978.

Augustauskas, R., & Lipnickas, A. (2020). Improved pixel-level pavement-defect segmentation using a deep autoencoder. Sensors, 20(9), 2557.

Bhardwaj, N., Kaur, G., & Singh, P K. (2017, October 4). A Systematic Review on Image Enhancement Techniques. Springer Nature, 227-235.

Chakraborty, T., S, U R K., Naik, S M., Panja, M., & Manvitha, B. (2024, January 17). Ten years of generative adversarial nets (GANs): a survey of the state-of-the-art. IOP Publishing, 5(1), 011001-011001.

Coghlan, S., & Quinn, T P. (2023, May 10). Ethics of using artificial intelligence (AI) in veterinary medicine. Springer Nature.

DeVries, T., & Taylor, G W. (2017, January 1). Improved Regularization of Convolutional Neural Networks with Cutout. Cornell University.

Dosovitskiy, A., & Brox, T. (2016). Generating images with perceptual similarity metrics based on deep networks. Advances in neural information processing systems, 29.

Gupta, R., Sharma, A., & Kumar, A. (2020, January 1). Super-resolution using GANs for Medical Imaging. Elsevier BV, 173, 28-35.

Hernández-García, Á., & König, P. (2018, January 1). Further Advantages of Data Augmentation on Convolutional Neural Networks. Springer Science+Business Media, 95-103.

Hu, L., Qin, M., Zhang, F., Du, Z., & Liu, R. (2021). RSCNN: A CNN-Based Method to Enhance Low-Light Remote-Sensing Images. Remote Sensing, 13(1), 62.

Karlık, B.; Aydın, S.; Kılıç, İ. (1996). Comparing Artificial Neural Networks (ANN) Image Compression Technique with Different Image Compression Techniques. Math. Comput. Applications, 1, 165-171.

Karlik, B. (2000). Comparison of widely used compression techniques for medical image. Proceedings of the PRASA, 105-110.

Karlık, B. (2006 June 29). Medical image compression by using vector quantization neural network (VQNN). Neural Network World, 16 (4), 341-348.

Khan, S., Ullah, I., Khan, F., Lee, Y., & Ullah, S. (2023, April 14). Historical Text Image Enhancement Using Image Scaling and Generative Adversarial Networks. Multidisciplinary Digital Publishing Institute, 23(8), 4003-4003.

Khosla, C., & Saini, B S. (2020). Enhancing Performance of Deep Learning Models with different Data Augmentation Techniques: A Survey. nternational Conference on Intelligent Engineering and Management (ICIEM), London, UK, June 1, 2020, 79-85, doi: 10.1109/ICIEM48762.2020.9160048.

Kirsten, L N., Piccoli, R., & Ribani, R. (2021, August 16). Evaluating deep neural networks for image document enhancement. https://doi.org/10.48550/arXiv.2106.15286

Kocyigit, Y., Korurek, M., & Karlik, B. (1999). ECG Data Compression by Artificial Neural Networks. Proceeding of the ELECO, 99, 338-340.

Long, J D., Shelhamer, E., & Darrell, T. (2014, January 1). Fully Convolutional Networks for Semantic Segmentation. Cornell University.https://doi.org/10.48550/arXiv.1411.4038

Mushtaq, Z. B., Nasti, S. M., Verma, C., Raboaca, M. S., Kumar, N., & Nasti, S. J. (2022). Super-resolution for noisy images using convolutional neural networks. Mathematics, 10(5), 777.

Mustafa, W A., Yazid, H., Mustafa, W A., Jamlos, M A., Zunaidi, I., Razlan, Z M., & Shahriman, A B. (2019, June 1). Image Enhancement Based on Discrete Cosine Transforms (DCT) and Discrete Wavelet Transform (DWT): A Review. IOP Publishing, 557(1), 012027-012027.

O'Mahony, N., Campbell, S., Carvalho, A., Harapanahalli, S., Hernandez, G. V., Krpalkova, L., ... & Walsh, J. (2020). Deep learning vs. traditional computer vision. In Advances in Computer Vision: Proceedings of the 2019 Computer Vision Conference (CVC), Volume 1 1 (pp. 128-144). Springer International Publishing.

Peng, Y., Sonka, M., & Chen, D Z. (2023, January 1). U-Net v2: Rethinking the Skip Connections of U-Net for Medical Image Segmentation. Cornell University. https://doi.org/10.48550/arXiv.2311.17791

Samek, W., Wiegand, T., & Müller, K. (2017, January 1). Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models. Cornell University.

Taylor, L., & Nitschke, G. (2017, January 1). Improving Deep Learning using Generic Data Augmentation. Cornell University. https://doi.org/10.48550/arXiv.1708.06020

Tensmeyer, C., & Martinez, T. (2017, November 1). Analysis of Convolutional Neural Networks for Document Image Classification. https://doi.org/10.48550/arXiv.1708.03273

Thomas, S A. (2021, January 1). Enhanced Transfer Learning Through Medical Imaging and Patient Demographic Data Fusion. Cornell University. https://doi.org/10.48550/arXiv.2111.14388

Tian, T., Wan, J., Song, Q., & Wei, Z. (2019). Clustering single-cell RNA-seq data with a model-based deep learning approach. *Nature Machine Intelligence*, *1*(4), 191-198.

Zhang, Y., Tiňo, P., Leonardis, A., & Tang, K. (2021, August 24). A Survey on Neural Network Interpretability. Institute of Electrical and Electronics Engineers, 5(5), 726-742.