Modified Base Autoencoder and Variational Autoencoder for Denoising Images in CIFAR-10 and MNIST Datasets

Jenan A. Alhijaj[1], Rana J. AL-Sukeinee[2], Asaad A. Alhijaj[1], Noor M. Al-Moosawi[1], Raidah S. Khudeyer[1]
[1]Department of Computer Science, College of Computer Science and Information Technology, University of Basrah, Basrah, Iraq
[2]Department of Physics, College of Science, University of Basrah, Basrah, Iraq
E-mail: jenan.alkereem@uobasrah.edu.iq, rana.jabbar@uobasrah.edu.iq, asaad.abdulhassan@uobasrah.edu.iq, almoosawinoor2@gmail.com, raidah.khudayer@uobasrah.edu.iq

Keywords*:* PSNR, SSIM, denoising image, autoencoder, variational autoencoders

Received: June 9, 2025

With the increasing volume of digital images, we must increase the quality of images for accuracy and visible applications, and we need ways to reduce the image noise while keeping important features such as edges, corners, and sharp details. In recent years, deep learning algorithms have become more significant for solving image denoising problems because they can simulate complex image patterns. This paper compares the performance of modified base AutoEncoders (AEs) and Variational Autoencoders (VAEs) models for image denoising in CIFAR-10 for color images and MNIST for grayscale images datasets. Our proposed modification to base AE and VAE architectures consists of changes in the encoder and decoder layers of feature extraction and reconstruction abilities, resulting in improved denoising performance. To simulate real-world image damages, data preparation involved normalization and the injection of Gaussian noise (0.5 for MNIST and 0.5 for CIFAR-10). With batch normalization and UpSampling2D layers with sigmoid outputs, the encoder-decoder architecture guaranteed the accuracy of spatial reconstruction, while VAE combined MSE with KL divergence for latent regularization, and AE optimized MSE reconstruction loss. The two models' performance was evaluated using important essential metrics: the Structural Similarity (SSIM) and the Peak Signal to Noise Ratio (PSNR). In both datasets, the results indicate that the VAE model outperforms the AE model in terms of image quality. The CIFAR-10 color dataset was given an SSIM of 0.954 and a PSNR of 32.86 dB, whereas the MNIST grayscale dataset provided an SSIM of 0.951 and a PSNR of 24.44 dB to the modified VAE model. In addition, the CIRAF-10 dataset achieved an SSIM of 0.891 and a PSNR of 27.72 dB, whereas the MNIST dataset was given an SSIM of 0.883 and a PSNR of 29.06 dB from the AE model. This study addresses how AE and VAE architectures differ in denoising performance across dataset complexities and the principles for optimal model selection.

Povzetek: Spremenjeni modeli VAE in AE so bili primerjani pri odstranjevanju šuma iz slik, pri čemer je VAE na podatkovnih zbirkah MNIST in CIFAR-10 dosegel boljšo kakovost rekonstrukcije.

1 Introduction

The study of deep learning, a branch of machine learning, has brought about a new phase in the advancement of neural networks [2]. A crucial component of neural networks, an autoencoder efficiently compresses (encodes) input data to its most basic components before using this compressed representation to reconstruct (decode) the original input. The reduction of the features from high dimensions to low dimensions leads to an enhanced quality of the digital images [1]. Autoencoder has several applications, such as data compression and image denoising [3]. In addition, abnormality Detection in industrial operations. In addition, Autoencoders can be used in recommendation systems that can be used for personalized suggestions and potential representations of users [4].

Similar to classic autoencoders but using a probabilistic framework, variational autoencoders (VAEs) are a kind of generative model; it's use generative models to produce new data in the form of variants of the input data that the models are trained on [5]. Neural networks known as VAEs take in input data and transform it into its latent representation, which consists of hidden layers for the variables mean and variance. then, the original data is recreated by converting these two hidden layers into sampled latent vectors [6]. VAEs have various applications like Latent Variable Modelling, Data Augmentation, Image Generation, and Anomaly Detection. Moreover, VAEs can produce new data points enabled by their probabilistic structure, which is especially helpful for tasks like image denoising [7].

To reduce noise while maintaining important image information, image denoising is a basic task in computer