Enhancing Lung Cancer Diagnosis Accuracy through Autoencoder-Based Reconstruction of Computed Tomography (CT) Lung Images

Mohammad Amin Pirian[a], Iman Heidari[a], Toktam Khatibi[b,*], Mohammad Mehdi Sepehri[b]

[a]*Industrial Engineering, Tarbiat Modares University*
[b]*Faculty of Industrial and Systems Engineering, Tarbiat Modares University*

Abstract

Lung cancer is a major global cause of cancer-related deaths, emphasizing the importance of early detection through chest imaging. Accurate reconstruction of computed tomography (CT) lung images plays a crucial role in the diagnosis and treatment planning of lung cancer patients. However, noise in CT images poses a significant challenge, hindering the precise interpretation of internal tissue structures. Low-dose CT, with reduced radiation risks, has gained popularity. Nonetheless, inherent noise compromises image quality, potentially impacting diagnostic performance. Denoising autoencoder and unsupervised deep learning algorithms offer a promising solution. A dataset of CT images from patients suspected of lung cancer was categorized into four disease groups to evaluate different autoencoder models. Results showed that designed autoencoders effectively reduced noise, enhancing overall image quality. The semi-supervised autoencoder exhibited superior performance, preserving fine details and enhancing diagnostic information. This research underscores autoencoder models' potential in improving lung cancer diagnosis accuracy by reconstructing CT lung images, emphasizing the importance of noise reduction techniques in enhancing image quality and diagnostic performance.

Keywords: Deep Learning, Autoencoder, Computed tomography images reconstruction, Image quality enhancement

1-Introduction

The analysis of medical images about cancer poses significant challenges due to limitations in sample collection, along with various issues such as noise, incomplete annotation, data dispersion, and the high dimensionality of images (characterized by a large number of variables). Consequently, the development of integrated computing approaches to effectively handle such data remains an intricate task. In recent times, numerous machine learning methodologies have been put forth as potential solutions to tackle these complex datasets (Reel et al., 2021). Unsupervised learning techniques play a pivotal role in identifying latent patterns within complex data while effectively overcoming inherent challenges. Notably, among the unsupervised evaluation methods, neural network-based approaches, such as Autoencoders (AEs) and Variational Autoencoders (VAEs) (Kingma & Welling, 2013, Rezende et al., 2014), have demonstrated promising performance across diverse datasets and contexts, including cancer diseases (Simidjievski et al., 2019), bacterial infection (Deng et al., 2019), and the identification of healthy patient tissues (Christopher Heje Grønbech et al., 2020).

The autoencoder is a powerful neural network architecture that learns to extract a concise representation of the input data, gradually reducing the dimensionality through its layers. As the information is processed, it converges to a bottleneck layer, which captures the most intrinsic features of the input data. The reconstruction process then proceeds in the reverse direction to recreate the original data. Through this compression-decompression mechanism, the algorithm gains a more refined data representation, effectively capturing the inherent relationships between the data variables. As a result, downstream analyses can benefit from a more precise and accurate understanding of the data structure (Belkin & Niyogi, 2003).

The primary objective of this research article is to assess and compare the performance of various autoencoder reconstruction models using computed tomography (CT) images from lung cancer patients. The ultimate aim is to offer valuable insights into selecting the most suitable reconstruction technique that can enhance the diagnostic accuracy for lung cancer patients.

*Corresponding Author. Email: toktam.khatibi@modares.ac.ir

2.Artificial Intelligence Methods for Medicine

For several years, machine learning and deep learning techniques have been extensively employed in the realm of medical image recognition. These models can be broadly categorized into three main types: supervised, unsupervised, and semi-supervised learning approaches.

In supervised learning methods, algorithms are trained using "labeled" data, wherein the input data is paired with corresponding labels or outcomes. During the training process, the algorithm learns to infer a function that can classify data accurately or predict results based on the provided labeled examples. This type of approach is commonly used in tasks such as identifying diseases in medical images or predicting patient outcomes based on their clinical data (Cunningham et al., 2008).The primary objective of unsupervised learning is to discern underlying structures and features within a training dataset without the use of data labels or annotations. This type of learning is particularly useful when dealing with unlabeled data. Unsupervised methods encompass a range of techniques, including clustering algorithms such as hierarchical clustering and k-means clustering, as well as dimensionality reduction approaches like principal component analysis (PCA), t-distributed Stochastic Neighbor Embedding (t-SNE), and uniform manifold approximation and projection (UMAP) (Com & Hinton, 2008, McInnes et al., 2018).

Among the unsupervised methods, artificial neural networks (ANNs), particularly auto-encoders (AEs) and variational auto-encoders (VAEs), have emerged as highly efficient tools for addressing diverse medical challenges and integrating different types of medical data (Munir et al., 2019). These neural network-based methods excel at extracting meaningful representations from complex medical data and have shown promise in various applications, including medical image analysis, disease classification, and patient clustering.

3.The Basis of Autoencoder Algorithm and Its Types

Autoencoders are a prominent unsupervised learning method in neural networks that aim to learn a lower-dimensional feature representation from input data. The primary structure of autoencoders consists of an input layer, a hidden layer, and an output layer (Ng, n.d.) .The dimensionality reduction in autoencoders results in the model retaining high-variance features while discarding low-variance ones. To achieve this, autoencoders are often combined with denoising methods to remove unimportant variations. This is achieved through a loss function that measures the distance between the compressed data and the reconstructed data, with common loss functions being mean squared error and relative entropy (Eraslan et al., 2019).Since the introduction of autoencoders, various variants have been proposed to address limitations, enhance models, and make improvements. Notable examples include Denoising Autoencoders (DAEs), Sparse Autoencoders (SAEs), and the more recent Variational Autoencoders (VAEs). DAEs are a class of autoencoders that introduce corruptions into some fields of the input matrix and use the corrupted data as input. This type of autoencoder has found significant application in image-related tasks (Vincent et al., 2008).

VAEs, on the other hand, are probabilistic generative models that make assumptions about the distribution of hidden layer features. They learn the true distribution of input features by using a Bayesian approach to approximate the latent space defined by mean "μ" and standard deviation "σ" of the latent variables. This feature allows VAEs to calculate probability distributions and generate new data instances (Baird et al., 1988). The key difference between traditional autoencoders and VAEs lies in the continuous nature of the latent space in the latter. VAEs are scalable to large datasets and can handle intractable posterior distributions by fitting an approximate inference or detection model, utilizing a parameterized change lower bound estimator. Their adaptability and capability to handle non-linear behaviors make VAEs particularly suitable for complex data analysis, such as data compression and dimensionality reduction (Hawkins et al., 2014). Recent benchmark studies have indicated VAEs as the superior method for detecting cancer subtypes in comparison to other types of autoencoders (Pratella et al., 2021).

4.AE Application in Cancer

Autoencoders (AEs) have found valuable applications in cancer data analysis, with several tools adopting different strategies and goals in this field. Specifically, these methods focus on two main areas: predicting drug response (Ladislav Rampášek et al., 2019, Chiu et al., 2019) and cancer diagnosis and classification (Franco et al., 2021, Way & Greene, 2017).The classification of cancer plays a crucial role in selecting appropriate treatment methods based on the cancer subtype or facilitating early diagnosis. Accurate cancer staging is closely associated with the prognosis and survival of cancer patients. To address these tasks, various tools have been proposed. For instance, Thibault et al. introduced a VAE-based learning method that classifies cancer types by capturing tissue-specific patterns and successfully identifies high-grade serous ovarian cancer (Way & Greene, 2017).

By leveraging the power of AEs and VAEs, researchers have made notable strides in understanding cancer heterogeneity and discovering valuable insights for personalized medicine and improved patient outcomes. These models enable the extraction of meaningful representations from complex cancer data, facilitating accurate classification and prediction tasks.

5.Data Description and Preprocessing

Two primary categories are recognized in lung cancer: non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). Non-small cell lung cancer constitutes the majority, affecting approximately 80-85% of patients. This type of lung cancer is further divided into three

subgroups, the subgroups differ based on the type of cells that initiate cancer in the patient's lung. Despite their distinctions, these subgroups share similarities in their treatment approach, which is why they are grouped together.

On the other hand, small cell lung cancer is less common, accounting for about 15-20% of cases (Rodriguez-Canales et al., 2016). Given the greater significance of non-small cell lung cancer in this study, the focus has been primarily on this subgroup.

This study's dataset comprises CT scan images of the lung, specifically related to non-small cell lung cancer (Open Data Commons Open Database License (ODbL) V1.0 — Open Data Commons: Legal Tools for Open Data, 2019). As previously mentioned, non-small cell lung cancer encompasses three main categories. When considering one group as representing normal lung tissue, there are a total of four general groups. Figure 1 demonstrates an example image from each of these four groups.
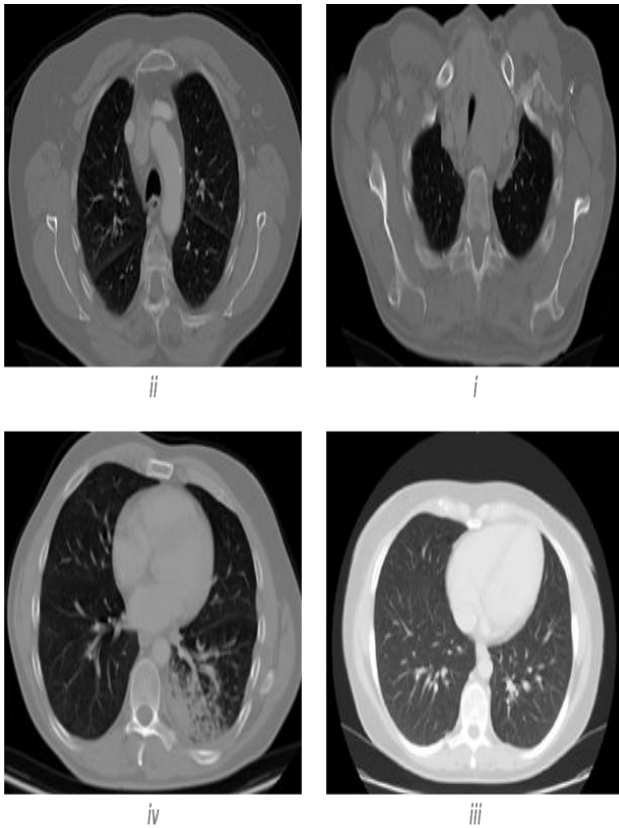

Fig 1. Examples of CT scan images of different types of non-small cell lung cancer

i) large cell carcinoma
ii) adenocarcinoma
iii) normal
iv) squamous cell carcinoma

5.1.Adding Noise to Images

In this study, noise has been intentionally introduced to the existing images to thoroughly evaluate the performance of the autoencoder models in handling noisy images. Among various methods of adding noise to images, the study utilizes point noise. Point noise involves adding random noise to each pixel of the image, where the noise values are drawn from a standard distribution. In this particular study, the variance of the standard distribution is set to (0.5).

To illustrate the effects of adding noise, Figure 2 displays examples of images from two classes before and after the noise addition process. The purpose of this experimentation is to examine how the autoencoder models can effectively denoise and reconstruct the original images despite the introduced noise. By subjecting the models to such tests, their robustness and ability to handle noisy data can be better assessed. The introduction of noise in image data is a crucial step in assessing the generalization and performance of autoencoders under realistic and noisy conditions, as real-world medical images often encounter various sources of noise during acquisition and transmission.

As depicted in Figure 2, black points have been deliberately added to all the images, effectively simulating the amount of noise introduced to each pixel of the image. These artificial noises have been intentionally incorporated into the images during this study to assess the performance of the models under conditions that resemble naturally occurring noise.
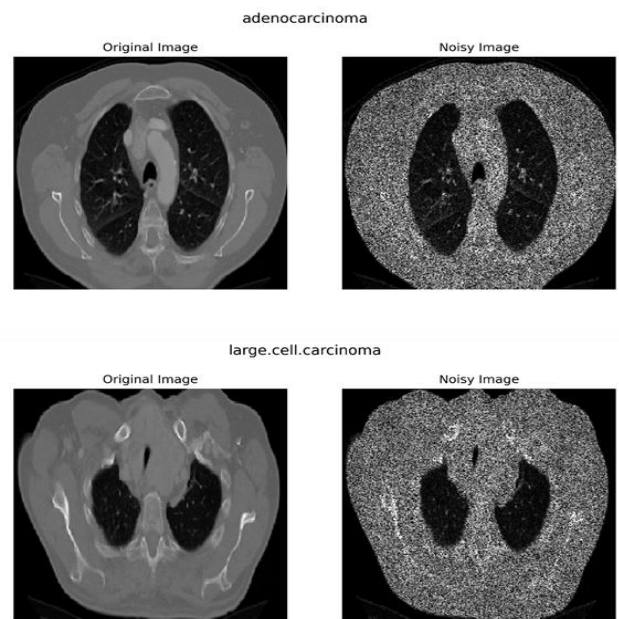

Fig 2 . Comparison of lung images before and after adding noise

6.Autoencoder Models

6.1 Classic Autoencoder Model

Indeed, measuring and evaluating the performance of the autoencoder model can be achieved by monitoring the loss function's value for each epoch during the training process. As depicted in Figure 3, the loss function

demonstrates a decreasing trend, indicating that the model's performance is improving as the number of epochs increases. The goal of the training process is to minimize the cost function, ultimately aiming to approach a value close to zero.

The loss function quantifies the discrepancy between the original input data and the reconstructed output generated by the autoencoder. The downward trend of the cost function signifies that the model is effectively learning to encode meaningful features and representations from the input data and subsequently reconstructing it with minimal error.
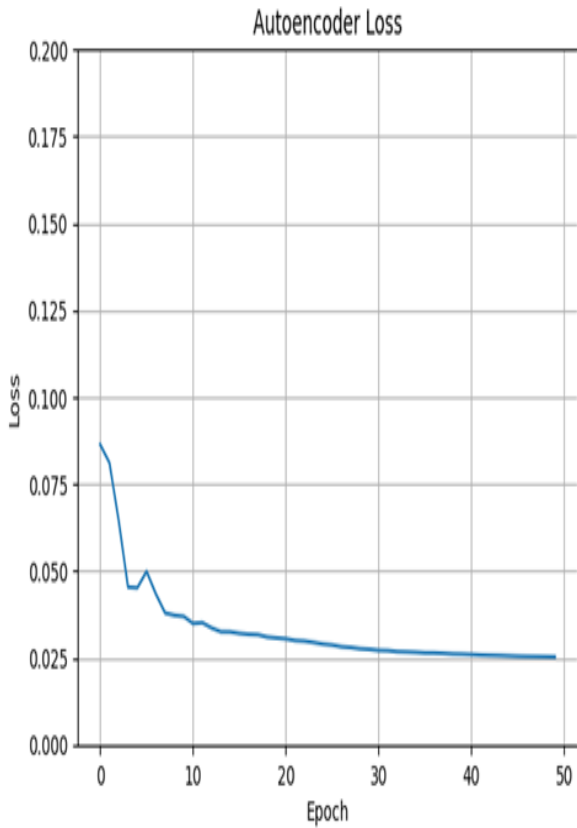
Fig 3 . loss function per epoch chart for the classic autoencoder model.

In the mentioned model, the loss function shows a good convergence to zero value, but this case alone cannot indicate that the model is perfect in removing noise. Because by looking at the output images of this model, it can be seen that the model does not work perfectly as expected from its loss function. Figure 4 illustrates the network architecture of this classic autoencoder.
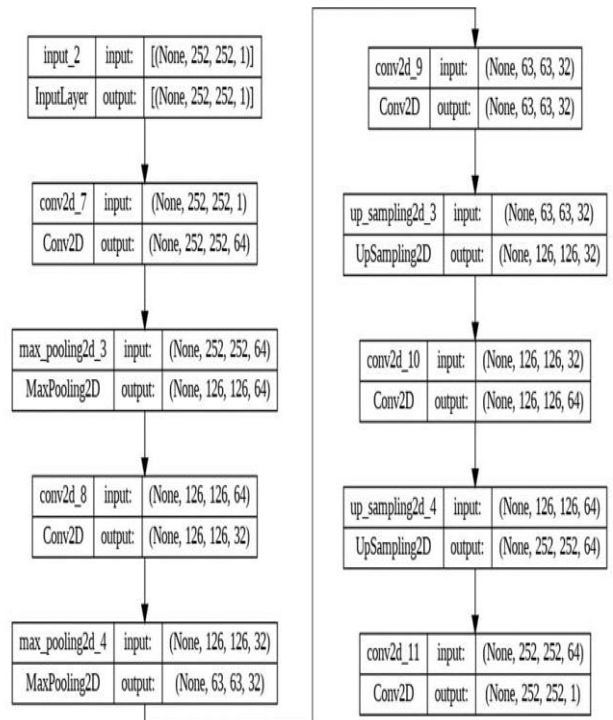
Fig 4 .The architecture of the constructed autoencoder network

6.2.Variational Autoencoder Model

This constructed model, being one of the types of autoencoder models, exhibits a notable difference from simple autoencoder models in terms of data representation within the hidden space. Specifically, the output of this model in the hidden layer comprises the parameters of a statistical distribution, unlike basic autoencoder models where the output is a vector.

The key distinction lies in the fact that, in this particular type of model, the hidden layer's output consists of both the mean and the logarithm of the variance. These parameters are indicative of a normal distribution function. By incorporating the logarithm of the variance, the model ensures that the predicted variance remains positive, as the logarithm of a negative value is undefined. In the cost function, this model follows two primary goals. The first objective is to enhance the quality of reconstructed images during the encoding and decoding processes. The second objective is to enforce the hidden space to conform to a normal distribution.

In the hidden layer of the modified autoencoder model, as shown in Figure 6, the compression of images during the encoding phase results in two outputs: the mean and the logarithm of the variance.
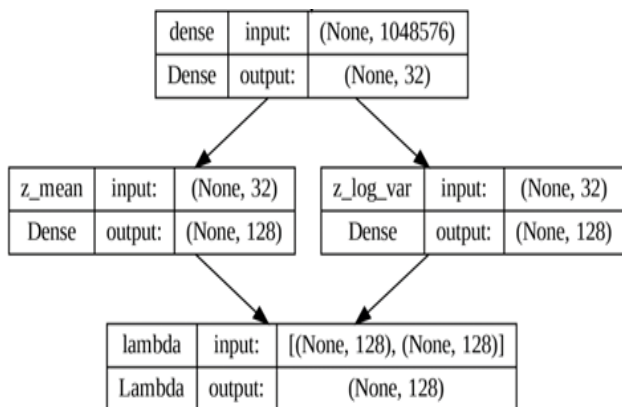
Fig 5 . The general structure of the hidden layer in the modified
VAE model

As depicted in Figure 5, the encoding process involves extracting both the mean and variance from the latent space. The mean and variance serve as essential parameters for a statistical distribution. During the decoding process, a random sample is drawn from this distribution to initiate the decoding and reconstruction of the data. The subsequent decoding steps follow a similar procedure as basic models. To evaluate the performance of the model during the training process, the loss chart for each epoch is utilized.
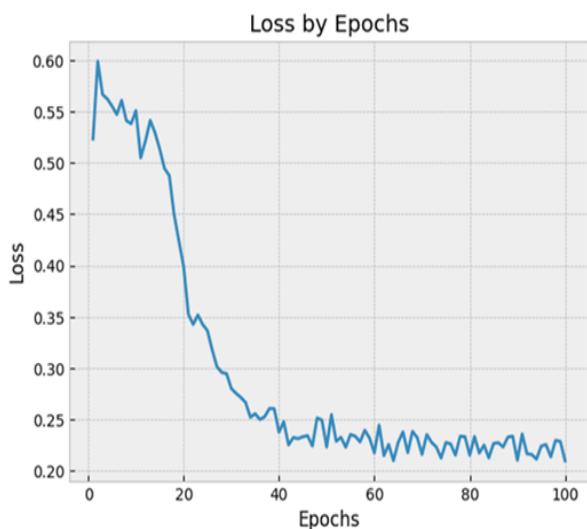
Fig 6 .Diagram of the loss function per epoch related to the

modified VAE model

Figure 6 displays the loss chart for the first 100 epochs, illustrating the trend of the loss function's value over the training iterations. As shown, the cost function exhibits a significant and steep downward slope around epoch 15, which indicates rapid improvement in the model's performance during the initial training phase. However, after epoch 42, the cost function starts to fluctuate around a value of approximately 0.23. This fluctuation suggests

that the model has reached a state of convergence, where the training progress has stabilized, and further training iterations are not leading to substantial improvements.

6.3.Semi-Supervised Autoencoder Model

A semi-supervised autoencoder (SSAE) is a specialized type of autoencoder that combines both unsupervised and supervised learning techniques. Unlike traditional autoencoders that rely solely on unsupervised learning from unlabeled data, SSAEs leverage a small set of labeled data in addition to a larger set of unlabeled data during training. This unique approach enables SSAEs to effectively remove noise from both the labeled and unlabeled data samples.

The encoder-decoder architecture of the SSAE learns relevant features from the labeled data and utilizes these learned features to denoise the unlabeled data. By incorporating labeled data, the SSAE can improve its denoising capabilities, leading to more accurate and reliable results.

Figure 8 illustrates the graph of the cost function for each epoch during the training of the SSAE model. It is important to note that the error function in this model is notably higher compared to the previous two models.
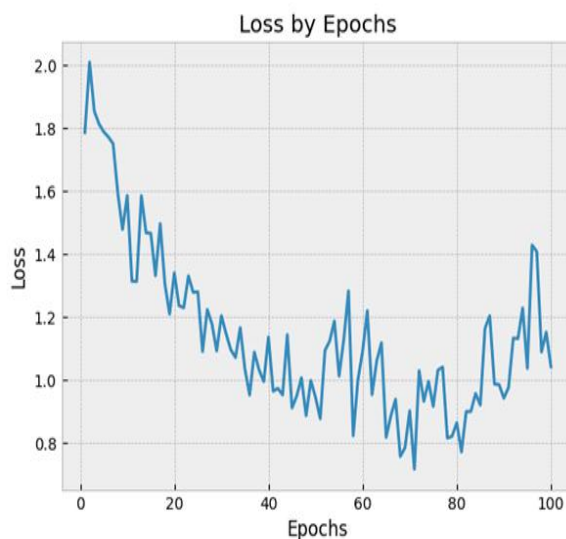
Fig 7 . Diagram of the loss function per epoch related to the semi-supervised autoencoder model

The higher error function in Figure 7 suggests that the denoising process in the SSAE model might be more challenging due to the use of labeled data. The model may be learning to generalize and denoise the unlabeled data effectively, but the additional complexity introduced by the labeled data might result in a higher overall cost function. However, the cost function alone may not accurately reflect the effectiveness of the model, as the reconstruction quality and denoising capabilities are equally important factors in assessing the model's performance.

As shown in Figure 8, the model successfully removes noise from the input images and provides reconstructed

images of acceptable quality. The noisy images are effectively denoised, and the model retains important features, resulting in high-quality reconstructions. The removal of noise from the images contributes to better accuracy in detection tasks.

The reconstructed images in Figure 8 indeed exhibit clear details and maintain the essential information present in the original images. This level of denoising and accurate reconstruction is critical for downstream applications, such as disease detection or image segmentation, where the quality and reliability of the reconstructed images play a significant role.

Fig 8. Examples of noisy images versus reconstructed images

In Figure 9, images with the highest loss function represent successful denoising, while Figure 10 shows challenging cases with the lowest loss function. Comparing images in Figures 9 and 10 reveals that the cost function is related to the color contrast of reconstructed images.
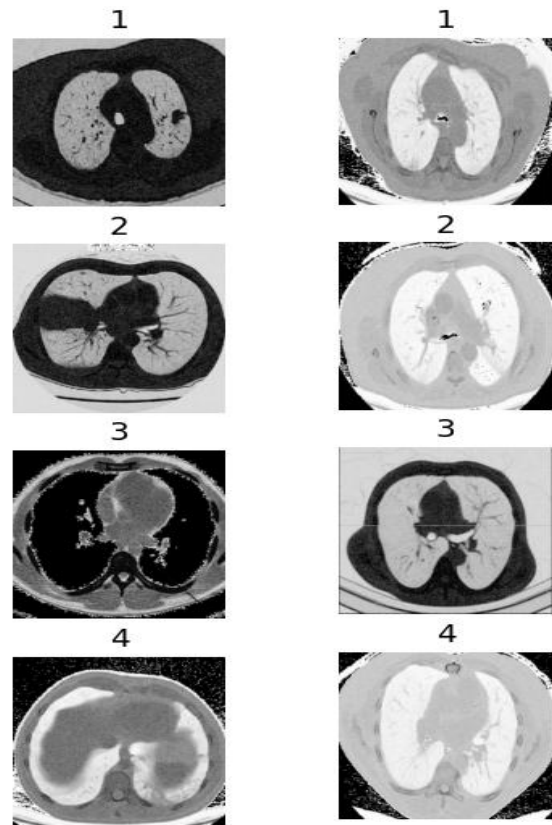
Fig 9 . Reconstructed images for each class with the highest amount of cost function

Fig 10. Reconstructed images for each class with the lowest cost function

7.Conclusion

This study demonstrates the successful application of an autoencoder-based approach to reconstructing noisy CT images in lung cancer patients. The trained autoencoder effectively reduced noise while preserving important anatomical details, thereby improving image quality. Quantitative evaluation criteria and qualitative analysis both confirmed the improvement of visual quality and diagnostic accuracy of reconstructed images. These findings highlight the potential of autoencoders to improve clinical decision-making and patient outcomes in lung cancer management and contribute to ongoing research on noise reduction in CT imaging.

Considering that the implementation of self-encryption algorithms in the field of CT images still has many issues to learn and develop, we believe that this research provides valuable insights on the selection of appropriate reconstruction techniques to increase the diagnostic accuracy of lung cancer patients and inspires progress. The future will be in the understanding and diagnosis of cancer.

References

Baird, P. A., Anderson, T. W., Newcombe, H. B., & Lowry, R. B. (1988). Genetic disorders in children and young adults: a population study. *American Journal of Human Genetics*, *42*(5), 677–693. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1715177

Belkin, M., & Niyogi, P. (2003). Laplacian Eigenmaps for Dimensionality Reduction and Data Representation. *Neural Computation*, *15*(6), 1373–1396. https://doi.org/10.1162/089976603321780317

Chiu, Y.-C., Chen, H.-I. H., Zhang, T., Zhang, S., Gorthi, A., Wang, L.-J., Huang, Y., & Chen, Y. (2019). Predicting drug response of tumors from integrated genomic profiles by deep neural networks. *BMC Medical Genomics*, *12*(S1). https://doi.org/10.1186/s12920-018-0460-9

Christopher Heje Grønbech, Maximillian Fornitz Vording, Pascal Timshel, Casper Kaae Sønderby, Pers, T. H., & Winther, O. (2020). *scVAE: variational auto-encoders for single-cell gene expression data*. *36*(16), 4415–4422. https://doi.org/10.1093/bioinformatics/btaa293

Com, L., & Hinton, G. (2008). Visualizing Data using t-SNE Laurens van der Maaten. *Journal of Machine Learning Research*, *9*, 2579–2605. https://www.jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf?fbcl

Contains information from Chest CT-Scan images Dataset, which is made available here under the Open Database License (*https://opendatacommons.org/licenses/odbl/1-0/*).

Cunningham, P., Cord, M., & Delany, S. J. (2008). Supervised Learning. *Machine Learning Techniques for Multimedia*, 21–49. https://doi.org/10.1007/978-3-540-75171-7_2

Deng, Y., Bao, F., Dai, Q., Wu, L. F., & Altschuler, S. J. (2019). Scalable analysis of cell-type composition from single-cell transcriptomics using deep recurrent learning. *Nature Methods*, *16*(4), 311–314. https://doi.org/10.1038/s41592-019-0353-7

Eraslan, G., Simon, L. M., Mircea, M., Mueller, N. S., & Theis, F. J. (2019). Single-cell RNA-seq denoising using a deep count autoencoder. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-018-07931-2

Franco, E. F., Rana, P., Aline Fernanda Cruz, Terry, M., Azevedo, V., Rommel, & Ghosh, P. (2021). *Performance Comparison of Deep Learning Autoencoders for Cancer Subtype Detection Using Multi-Omics Data*. *13*(9), 2013–2013. https://doi.org/10.3390/cancers13092013

Hawkins, S. H., Korecki, J. N., Balagurunathan, Y., Yuhua Gu, Kumar, V., Basu, S., Hall, L. O., Goldgof, D. B., Gatenby, R. A., & Gillies, R. J. (2014). Predicting Outcomes of Nonsmall Cell Lung Cancer Using CT Image Features. *IEEE Access*, *2*, 1418–1426. https://doi.org/10.1109/access.2014.2373335

Kingma, D. P., & Welling, M. (2013). Auto-Encoding Variational Bayes. *ArXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1312.6114

Ladislav Rampášek, Hidru, D., Smirnov, P. A., Haibe-Kains, B., & Goldenberg, A. (2019). *Dr.VAE: improving drug response prediction via modeling of drug perturbation effects*. *35*(19), 3743–3751. https://doi.org/10.1093/bioinformatics/btz158

McInnes, L., Healy, J., & Melville, J. (2018). *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction*. ArXiv.org. https://arxiv.org/abs/1802.03426

Munir, K., Elahi, H., Ayub, A., Frezza, F., & Rizzi, A. (2019). Cancer Diagnosis Using Deep Learning: A Bibliographic Review. *Cancers*, *11*(9), 1235. https://doi.org/10.3390/cancers11091235

Ng, A. (n.d.). *CS294A Lecture notes Sparse autoencoder*. https://graphics.stanford.edu/courses/cs233-21-spring/ReferencedPapers/SAE.pdf

Pratella, D., Ait-El-Mkadem Saadi, S., Bannwarth, S., Paquis-Fluckinger, V., & Bottini, S. (2021). A Survey of Autoencoder Algorithms to Pave the Diagnosis of Rare Diseases. *International Journal of Molecular Sciences*, *22*(19), 10891. https://doi.org/10.3390/ijms221910891

Reel, P. S., Reel, S., Pearson, E., Trucco, E., & Jefferson, E. (2021). Using machine learning approaches for multi-omics data analysis: A review. *Biotechnology Advances*, *49*, 107739. https://doi.org/10.1016/j.biotechadv.2021.107739

Rezende, D. J., Mohamed, S., & Wierstra, D. (2014, May 30). *Stochastic Backpropagation and Approximate Inference in Deep Generative Models*. ArXiv.org. https://doi.org/10.48550/arXiv.1401.4082

Rodriguez-Canales, J., Parra-Cuentas, E., & Wistuba, I. I. (2016). Diagnosis and Molecular Classification of Lung Cancer. *Cancer Treatment and Research*, *170*, 25–46. https://doi.org/10.1007/978-3-319-40389-2_2

Simidjievski, N., Bodnar, C., Tariq, I., Scherer, P., Andres Terre, H., Shams, Z., Jamnik, M., & Liò, P. (2019). Variational Autoencoders for Cancer Data Integration: Design Principles and Computational Practice. *Frontiers in Genetics*, *10*. https://doi.org/10.3389/fgene.2019.01205

Vincent, P., Larochelle, H., Bengio, Y., & Manzagol, P.-A. (2008). Extracting and composing robust features with denoising autoencoders. *Proceedings of the 25th International Conference on Machine Learning - ICML '08*. https://doi.org/10.1145/1390156.1390294

Way, G. P., & Greene, C. S. (2017). Extracting a biologically relevant latent space from cancer transcriptomes with variational autoencoders. *Biocomputing 2018*. https://doi.org/10.1142/9789813235533_0008