Deep Learning-Based Generative AI for Segmenting Necrotic Lung Lesions in CT Images Using Self-Supervised Contrastive Learning

G. Indumathi, Sudharsan T. S., Ashok I. and Tharunkumar M.
Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India

Keywords: Convolutional Neural Networks (CNNs), Necrotic Lung Lesions, CT Images, Pulmonary Diseases, Segmentation, Variational Autoencoders (VAEs), Self-Supervised Learning, Contrastive Learning, Feature Representation, Annotated Datasets, Generalization, Lesion Morphology, Medical Image Processing, Clinical Applications.

Abstract: Segmenting necrotic lung lesions in CT images plays a vital role in diagnosing and managing pulmonary diseases. However, traditional methods often struggle with the complex shapes and varying appearances of lesions while relying heavily on manually annotated datasets. To overcome these limitations, we introduce an innovative framework that combines Variational Autoencoders (VAEs) and self-supervised contrastive learning for more accurate and efficient segmentation. The VAE helps the model learn compact and meaningful representations of CT images, while contrastive pretraining enhances these features using unlabeled data, improving generalization across different datasets. This approach not only reduces dependency on manual annotations but also excels in capturing fine lesion boundaries and handling diverse lesion appearances. By advancing medical image segmentation, our method provides a robust, scalable, and efficient solution to key clinical challenges, ultimately aiding in early detection and treatment planning for lung diseases.

1 INTRODUCTION

Medical imaging, particularly computed tomography (CT), plays a crucial role in diagnosing and managing lung diseases. Identifying necrotic lung lesions which may be associated with infections, cancer, or inflammatory condition is essential for effective treatment. However, manually detecting and segmenting these lesions is challenging, time-consuming, and often subjective, varying from one radiologist to another. This highlights the urgent need for automated and reliable segmentation methods to improve accuracy and efficiency in clinical settings.
With recent advancements in deep learning, medical image analysis has seen remarkable improvements. Generative models like Variational Autoencoders (VAEs) can effectively capture meaningful patterns in complex medical data while preserving critical diagnostic details. When combined with self-supervised contrastive learning, these models become even more powerful by learning from 1odelling1 medical images.

In this study, we present an innovative approach that integrates VAE-based generative 1odelling with self-supervised contrastive learning to segment necrotic lung lesions in CT scans. The VAE extracts compact, informative features, while contrastive learning refines them by distinguishing between similar and different patterns. This synergy improves segmentation accuracy and enables the model to generalize across diverse lung lesion types. By automating lesion detection, our method provides a scalable and efficient solution that enhances clinical decision-making, ultimately leading to better patient outcomes.

2 LITERATURE SURVEY

Medical image segmentation plays a vital role in diagnosing and managing pulmonary diseases, especially when identifying necrotic lung lesions in CT scans. These lesions often have irregular shapes and varying appearances, making them difficult to segment using traditional techniques like thresholding or region-growing methods. As a result,

Indumathi, G., S., S. T., I., A. and M., T.
Deep Learning-Based Generative AI for Segmenting Necrotic Lung Lesions in CT Images Using Self-Supervised Contrastive Learning.
DOI: 10.5220/0013932200004919

In *Proceedings of the 1st International Conference on Research and Development in Information, Communication, and Computing Technologies (ICRDICCT'25 2025) - Volume 5*, pages 512-518
ISBN: 978-989-758-777-1

deep learning-based generative models have emerged as more accurate and robust solutions for this task.

Deep learning architectures like Convolutional Neural Networks (CNNs) and U-Net have shown remarkable success in medical image segmentation. However, these methods often require large annotated datasets, which can be a major limitation due to the time and expertise needed for manual labeling. To overcome this, Variational Autoencoders (VAEs) first introduced by Kingma and Welling (2013) offer an effective alternative by learning compact, meaningful representations of high-dimensional data. This capability makes them valuable for both segmentation and data augmentation in medical imaging.

To further reduce reliance on labeled datasets, Self-Supervised Learning (SSL) enables models to learn from unlabeled data through pretext tasks like image reconstruction. Contrastive learning, a subset of SSL, enhances feature extraction by ensuring that similar images (positive pairs) are grouped together while distinct images (negative pairs) are kept apart. Popular contrastive learning methods like SimCLR (Chen et al., 2020) and MoCo (He et al., 2020) have significantly improved the generalization and robustness of deep learning models in medical image analysis.

By combining VAEs with self-supervised contrastive learning, we can develop a powerful and efficient solution for necrotic lung lesion segmentation. This approach reduces the need for manual annotation, enhances segmentation accuracy, and allows the model to adapt to diverse imaging conditions. Moving forward, refining contrastive loss functions and exploring hybrid deep learning models could further improve segmentation performance, making AI-driven medical imaging even more reliable and effective for clinical applications.

3 PROBLEM STATEMENT

Segmenting necrotic lung lesions in CT scans is a complex task due to their irregular shapes, low contrast, and limited availability of annotated data. Traditional deep learning models rely heavily on manual labeling, which is time-consuming and labor-intensive. To overcome these challenges, this research introduces a generative AI framework that combines Variational Autoencoders (VAEs) with self-supervised contrastive learning. This approach enhances segmentation accuracy, reduces the need for manual annotations, and improves the model’s ability to generalize across diverse datasets. By leveraging unlabeled data and learning meaningful patterns, this method offers a more efficient and scalable solution for medical image analysis.

3.1 Thresholding Method

Thresholding is a basic technique for image segmentation. It chooses a threshold value and separates pixels into two classes: those that are above the threshold and those that are below. This technique can be applied to separate the objects of interest, such as potential lung lesion regions, from the background in an image.

How it applies to lung lesions identification

This method combines VAEs, contrastive learning, and U-Net to enhance necrotic lung lesion segmentation in CT images, achieving high accuracy while minimizing the need for manual annotations.

Limitations

The system relies on high-quality CT scans, requires extensive computational resources, may struggle with extreme lesion variations, and needs clinical validation.

3.2 K-Means Clustering

K-means clustering is an unsupervised machinelearning algorithm. It attempts to partition 'n' observations into 'k' clusters in which each observation belongs to the cluster-1 with the nearest mean, serving as a prototype of the cluster-2.

Manual Analysis:

This is the conventional method in which a trained expert (such as a pulmonologist) visually studies the medical images (MRI, CT scans) and manually marks the region of the tumor.

3.2.1 Disadvantages

1. **High Computational Cost:** Training VAEs and contrastive learning models demands high GPU resources, making real-time clinical deployment costly and technically challenging for widespread medical use.
2. **Limited Generalization:** The model's reliability may be affected by unseen data due to variations in lesion shape, CT scan quality, and diverse patient demographics.
3. **Interpretability Issues:** Generative AI models operate like black boxes, making it hard for clinicians to interpret predictions, which limits trust in medical applications.
4. **Clinical Validation Requirement:** Extensive

real-world validation is essential to confirm accuracy, robustness, and compliance with medical standards before this approach can be widely adopted.

4 PROPOSED SYSTEM

Our system combines Variational Autoencoders (VAEs) and Self-Supervised Contrastive Learning to segment necrotic lung lesions in CT images. The figure1 shows the Block diagram of the proposed system VAEs extract meaningful latent features, while contrastive learning refines these features using unlabeled data. This method enhances segmentation accuracy, minimizes reliance on manual annotations, and improves generalization across different datasets, making it a more efficient and scalable solution for medical imaging.

4.1 Variational Autoencoders (VAEs)

1. **Learning Data Distribution:** VAEs are generative models. They are not simply learning how to classify images; they learn the underlying probability distribution of data. Think of learning the "essence" of what makes a brain image be a brain image and what variations are possible.
2. **Image Enhancement:** This learned knowledge can be used to manipulate and improve images.
3. **Noise Removal:** VAEs can reconstruct images to remove noise and artifacts, which could blur important information, making images clearer for further processing.
4. **Detailed Representation:** VAEs can also produce a more detailed or improved representation of images. It can be enhanced through the feature representation by drawing out subtle features and filling in gaps, which is easier to see abnormalities.
5. **Improved Classification Accuracy:** Through the enhancement of the quality and clarity of the input images, VAEs indirectly help to improve the performance of downstream tasks such as diseaseclassification. The clearer the images are, the easier it is for a model trying to detect a tumor.

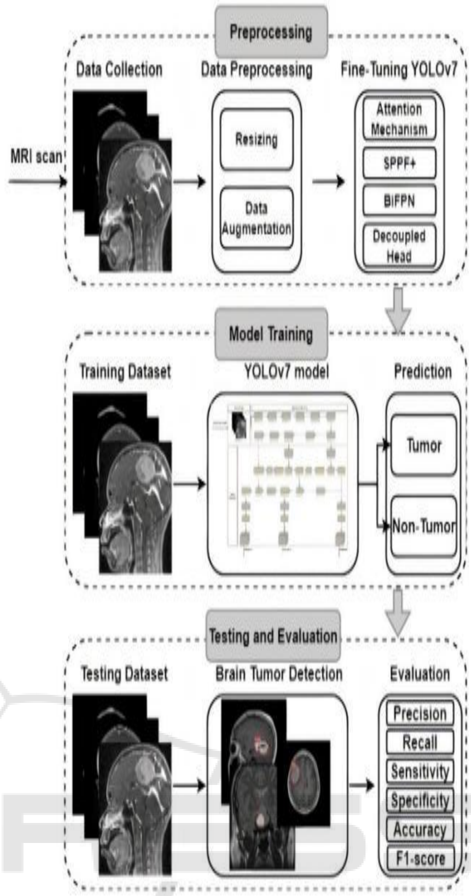

Figure 1: Block Diagram of the Proposed System.

4.2 Convolutional Neural Networks (CNNs)

1. **Image Classification:** Image classification is arguably the application field where CNNs are most effectively used. Spatial hierarchies of features for images are where they excel particularly well. Therefore, in case of brain tumors, they learned to classify one category of the brain tissue different from another.
2. **Labeled Dataset:** This is a kind of supervised learning model of CNNs, so it requires data to be labeled. The paragraph refers to the dataset of the labeled brain MRI or CT images, for instance, benign tumor, malignant tumor, and normal tissue.
3. **Feature Extraction:**The CNN learns to automatically extract the features relevant from images. Features are not pre-defined by humans, but instead are learned by the network. In the case of brain tumor detection, the shape, size, texture, and other features about

the tumor would be examples.

4. **Automation and Efficiency:** The most important advantage of CNNs is that they automate the process of feature extraction and classification. This makes the detection of brain tumors faster, more objective, and potentially more accurate than manual analysis.
5. **Early Diagnosis and Better Outcomes:** By enabling rapid and precise identification of brain tumors, CNNs contribute to earlier diagnosis, which is crucial for improving treatment outcomes and patient survival.

4.3 Advantages

1. **Reduced Annotation Dependency:** By learning from unlabeled CT images, self-supervised contrastive learning reduces the reliance on large annotated datasets, making model training more efficient.
2. **Improved Segmentation Accuracy:** Combining Variational Autoencoders (VAEs) with contrastive learning improves feature extraction, resulting in more accurate detection of lesion boundaries in CT images.
3. **Better Generalization:** The model effectively adapts to diverse datasets, ensuring greater robustness across various imaging conditions and patient populations for more reliable segmentation..
4. **Enhanced Feature Representation:** VAEs extract meaningful latent features, effectively capturing variations in lesion appearance and structure for more accurate and reliable segmentation.
5. **Efficient Handling of Complex Morphologies:** The framework accurately segments lesions with irregular shapes, low contrast, and diverse textures, offering superior performance compared to traditional methods.
6. **Scalability:** After training, the system can process large datasets without extra annotations, making it highly efficient and practical for clinical applications.
7. **Reduced Human Effort and Time:** Automating lesion segmentation accelerates diagnosis and treatment planning, reducing radiologists' workload and improving efficiency in medical imaging analysis.
8. **Robust Against Label Noise:** Self-supervised learning reduces errors caused by inconsistent manual annotations, enhancing model reliability and ensuring more accurate and consistent lesion segmentation.
9. **Potential for Transfer Learning:** The learned representations are highly adaptable, allowing fine-tuning for other medical imaging tasks, improving versatility and expanding clinical applications.
10. **Supports Early Disease Detection:** Enhanced segmentation enables early detection of necrotic lung conditions, facilitating timely interventions and improving patient outcomes through more accurate diagnoses and treatment planning.

5 MODULE DESCRIPTION

5.1 Convolutional Neural Networks (CNNs)

Focus on CNNs as the primary architecture, with its ability to excel in image analysis and feature extraction.

Specific CNN **Architectures:** Discuss specific architectures of CNNs used, such as ResNet, VGG, Inception, and the reason for using them.

1. **Hybrid Models:** If applicable, discuss hybrid models that combine CNNs with other deep learning architectures, such as Recurrent Neural Networks or Transformers.

5.2 Variational Autoencoder (VAE)

1. **VAE for Data Augmentation:** The VAE learns the underlying distribution of healthy brain images. It can then generate synthetic, but realistic, healthy images to augment the training data, thereby improving the robustness of the model and reducing overfitting, especially in cases where the data is scarce.
2. **VAE for Feature Extraction:** The VAE's latent space (the compressed representation) can be used to extract features from the MRI images. These features, combined with the CNN's features, can improve the accuracy of tumor classification.
3. **VAE for Anomaly Detection:** Train the VAE on the healthy brain images to learn their normal representation. Tumor containing images will probably have a high reconstruction error while passing through the VAE since they are anomalous and can be used for potential tumor detection.

4. **Hybrid Model:** It should be explicitly stated if the model is a hybrid, which case it combines CNNs ,VAEs in some way, followed by the interaction between the two components.

5.3 Data Preprocessing and Augmentation

1. **Standard Preprocessing:** Outline the standard preprocessing steps for images (noise reduction, contrast enhancement, etc.).
2. **VAE-Based Augmentation:** Explain how the VAE is used to generate synthetic data. Mention specific techniques used to ensure the generated images are realistic and diverse.

5.4 Training and Evaluation

1. **Training the VAE:** Describe the training process for the VAE, including the loss function (typically a combination of reconstruction loss and KL divergence) and optimization techniques.
2. **Training the CNN (with or without VAE features):** Describe how one trains the CNN, potentially making use of VAE-extracted features.
3. **Evaluation Metrics:** State standard evaluation metrics (accuracy, sensitivity, specificity, F1-score, AUC). If VAE is used for anomaly detection, discuss the relevant metrics.

5.5 Potential Applications

Figure 2: : Sample Lung Cancer Image.

Clinical Diagnosis and CAD: Discuss applications in the clinical field, computer-aided diagnosis

5.6 Limitations

1. **VAE Training Complexity:** Recognize the challenge in training VAEs effectively.
2. **Interpretability:** Explain the difficulties in interpreting the latent space of the VAE and its contribution to tumor detection.the figure 2 shows the sample lung cancer image .

6 IMPLEMENTATION

Segmenting necrotic lung lesions in CT images is challenging, but deep learning-based generative AI enhances both accuracy and efficiency. Our method follows a structured pipeline, integrating data preprocessing, model training, self-supervised contrastive learning, and final segmentation to achieve highly precise and reliable results.

6.1 Preparing the Data

Before training an AI model, obtaining high-quality CT scan datasets is essential. We use publicly available datasets like LIDC-IDRI or hospital-acquired medical images. Raw CT scans often contain noise and irrelevant details that can affect model accuracy, so we apply preprocessing techniques to improve data quality. Standardization normalizes pixel intensities for consistency, while lung isolation removes unnecessary regions using thresholding and morphological operations. To enhance adaptability, data augmentation introduces variations like rotation, flipping, and contrast adjustments. Since 3D CT scans are computationally heavy, patch extraction breaks them into smaller 2D sections, preserving lesion details. These steps optimize segmentation accuracy and model performance.

6.2 Learning Features with Variational Autoencoders

Since medical image datasets are often limited, Variational Autoencoders (VAEs) help the model recognize meaningful patterns and generate realistic representations of CT images. VAEs consist of three key parts. First, the encoder compresses CT scans into a compact feature space, capturing important anatomical details. Instead of simply memorizing images, latent space sampling allows the model to learn underlying patterns, making it more adaptable to new data. Finally, the decoder reconstructs images from this compressed information while preserving critical lesion details. To ensure accuracy, we use Reconstruction Loss to keep generated images close to the originals and KL Divergence Loss to prevent overfitting and maintain a well-structured feature space.

6.3 Enhancing Feature Learning with Self-Supervised Contrastive Learning

Since labeling medical images is time-consuming and requires expert annotation, contrastive learning helps the model differentiate between similar and different lung lesions without relying on manual labels. The process begins by creating positive pairs (two slightly modified versions of the same CT scan) and negative pairs (images from different patients). These pairs are then processed through a Siamese Network or SimCLR framework, which teaches the model to pull similar images closer in the feature space while pushing dissimilar ones apart. To refine this learning, we use Contrastive Loss (NT-Xent Loss), which strengthens feature representations. This approach enhances the model's ability to accurately identify lung lesions, even in cases where labeled data is scarce, improving segmentation performance and generalization.

6.4 Performing Segmentation with U-Net

For the final segmentation step, we integrate the learned features into a U-Net model, a widely used deep learning architecture for medical image segmentation. The encoder (backbone) is initialized with contrastive learning-pretrained weights, allowing it to extract meaningful high-level features from CT scans. The decoder then reconstructs a pixel-wise segmentation mask, accurately outlining necrotic lung lesions. To preserve fine details, skip connections are used, ensuring that important spatial information from the encoder is retained throughout the network. To optimize segmentation accuracy, we employ a combination of Dice Loss, which measures overlap accuracy between the predicted mask and actual lesion, and Binary Cross-Entropy (BCE) Loss, which ensures accurate classification of lesion and non-lesion regions, improving overall model performance.

6.5 Evaluating Performance and Testing on New Data

To ensure our model performs well across various datasets, we evaluate it using key performance metrics. The Dice Similarity Coefficient (DSC) measures how accurately the predicted lesion mask overlaps with the actual ground truth, ensuring precise segmentation. Intersection over Union (IoU) further assesses segmentation accuracy by comparing the predicted and actual lesion areas. Additionally, we analyze precision, recall, and F1-score to detect false positives and false negatives, ensuring reliable performance. To confirm the model's robustness, we test it on unseen external datasets, verifying its ability to generalize across different imaging conditions, scanner types, and patient variations for real-world applicability.

7 CONCLUSIONS

The proposed deep learning-based generative AI approach provides an effective solution for segmenting necrotic lung lesions in CT images by integrating Variational Autoencoders (VAEs), self-supervised contrastive learning, and U-Net segmentation. VAEs help extract meaningful features from medical images, enabling the model to learn rich representations of lung structures. Contrastive learning, on the other hand, strengthens the model's ability to differentiate between healthy and diseased lung tissue without relying on extensive manually labeled data. This significantly enhances the model's generalization capability, making it more adaptable to real-world clinical applications.

By incorporating U-Net, which is well-suited for medical image segmentation, the framework ensures highly accurate, pixel-level detection of necrotic lung lesions. The encoder, pretrained with contrastive learning, extracts high-level features, while the decoder reconstructs precise segmentation masks. To validate performance, we use key evaluation metrics such as Dice Similarity Coefficient (DSC), Intersection over Union (IoU), precision, and recall, which confirm the model's reliability.

By reducing dependence on large annotated datasets, this scalable, automated, and highly accurate approach offers an advanced tool for early diagnosis and treatment planning, ultimately improving healthcare outcomes for lung disease patients.

REFERENCES

C. Peng, X. Zhang, G. Yu, G. Luo, and J. Sun, ''Large kernel matters Improve semantic segmentation by global convolutional network,'' inProc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017,pp. 4353–4361.

G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi,M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, and I. C. Sanchez,''A survey on deep learning in medical image analysis,'' Med. Image Anal.,vol. 42, pp. 60–88, Dec. 2017.

G. Pezzano, V. R. Ripoll, and P. Radeva, ''CoLe-CNN: Context-learningconvolutional neural network with adaptive loss function for lungnodulesegmentation,'' Comput. Methods Programs Biomed., vol. 198, Jan. 2021,Art. no. 105792.

H. Cao, H. Liu, E. Song, C.-C. Hung, G. Ma, X. Xu, R. Jin, and J. Lu,''Dual-branch residual network for lung nodule segmentation,'' Appl. SoftComput., vol. 86, Jan. 2020, Art. no. 105934.

O. Ronneberger, P. Fischer, and T. Brox, ''UNet: Convolutional networksfor biomedical image segmentation,'' in Proc. Int. Conf. Med. ImageComput. Comput.-Assist. Intervent. Cham, Switzerland: Springer, 2015,pp. 234–241.

S. B. Knight, P. A. Crosbie, H. Balata, J. Chudziak, T. Hussell, and C. Dive,''Progress and prospects of early detection in lung cancer,'' Open Biol.,vol. 7, no. 9, Sep. 2017, Art. no. 170070.

S. Wang, M. Zhou, Z. Liu, Z. Liu, D. Gu, Y. Zang, D. Dong, O. Gevaert,and J. Tian, ''Central focused convolutional neural networks: Developinga data-driven model for lung nodule segmentation,'' Med. Image Anal.,vol. 40, pp. 172–183, Aug. 2017.

S. Luo, J. Zhang, N. Xiao, Y. Qiang, K. Li, J. Zhao, L. Meng, and P. Song,''DAS-Net: A lung nodule segmentation method based on adaptivedualbranchatte ntion and shadow mapping,'' Appl. Intell., vol. 52, no. 13,pp. 15617–15631, 2022.

V. Badrinarayanan, A. Handa, and R. Cipolla, ''SegNet: A deep convolutionalencoder–decoder architecture for robust semantic pixel-wiselabelling,'' 2015, arXiv:1505.07293.

X. Dong, S. Xu, Y. Liu, A.Wang, M. I. Saripan, L. Li, X. Zhang, and L. Lu,''Multi-view secondary input collaborative deep learning for lung nodule3D segmentation,'' Cancer Imag., vol. 20, no. 1, pp. 1–13, Dec. 2020.