MV-CT-Net: Exploiting multi-view context fusion for enhanced chest CT image analysis

Axel Bessy[1,2], Thomas Barba[3], Hamid Ladjal[2], Antoine Richard[3], Mathieu Lefort[2], Simon Achard[1], and Alexandre Meyer[2]

[1] ATOS, Villeurbanne, France
[2] LIRIS, UMR5205, CNRS, INSA de Lyon, Université Claude Bernard Lyon 1, Villeurbanne, France
[3] Hospices Civils de Lyon (HCL), Lyon, France

Abstract. Automated medical image analysis has become an essential tool for computer-aided diagnosis, supporting accurate and large-scale detection of multiple pathologies across diverse clinical applications. In this work, we propose **MV-CT-Net**, a multi-view deep learning model for detecting multiple thoracic pathologies from 3D chest CT scans. The model uses axial, coronal, and sagittal views together to capture more complete anatomical information. To handle the imbalance between common and rare diseases, we introduce a class-specific threshold strategy that improves recall and F1 score. Results on the CT-RATE benchmark show that MV-CT-Net achieves performance comparable to existing state-of-the-art methods across several evaluation metrics while detecting a wider range of thoracic pathologies.

Keywords: Medical Image Analysis · Chest CT Analysis · Thoracic Pathology Detection · Multi-Label Classification

1 Introduction

Automated analysis of medical images has become a cornerstone of computer-aided diagnosis, spanning domains as diverse as dermatology, cardiology, pulmonology, and gastroenterology. Each clinical domain poses distinct challenges in terms of image acquisition, annotation, and pathology manifestation, making the development of general-purpose diagnostic systems particularly difficult. Because medical data are inherently sensitive, publicly available datasets tend to be small and narrowly scoped, which has led to a proliferation of highly specialized models, each targeting a limited set of conditions within a single domain [6,12].

While these specialized approaches can achieve strong performance on their respective tasks, they suffer from a fundamental limitation: in clinical practice, different specialists may interpret the same image through the lens of their own expertise, potentially overlooking abnormalities outside their field. Given the cost of repeated examinations and referrals, a unified model that simultaneously screens for a broad spectrum of pathologies could facilitate earlier triage and more comprehensive diagnoses.

Recent efforts to curate large-scale, multi-label datasets for thoracic computed tomography (CT) scans [4] have opened new possibilities for training such unified models. In this work, we propose **MV-CT-Net**, a multi-view architecture that jointly processes axial, coronal, and sagittal planes of 3D chest CT volumes jointly, aiming to match or surpass the accuracy of single-view specialist models while covering a wider range of thoracic abnormalities.

Our key contributions are summarized as follows.

1. We propose MV-CT-Net, a multi-view fusion architecture that leverages complementary anatomical information from axial, coronal, and sagittal planes of chest CT volumes for multi-label thoracic pathology classification.
2. We introduce a class-specific threshold learning strategy to mitigate severe class imbalance, leading to improved recall and F1-score performance, particularly for rare pathologies.
3. We demonstrate competitive or superior performance compared with state-of-the-art methods on the CT-RATE benchmark metrics while covering a broader range of thoracic pathologies.

2 Related works

Recent advances in deep learning for chest CT analysis mainly follow two directions: supervised visual encoder architectures for task-specific volumetric representation learning, primarily designed for downstream tasks such as classification, and large scale foundation models trained using self-supervised or vision language pre-training to learn generalizable and transferable representations from large scale imaging or multimodal datasets.

Visual Encoders Multi-label classification of 3D chest CT volumes has attracted growing interest, with several visual encoder architectures proposed in recent years. CT-Net [2] pioneered the use of a 2D CNN backbone for volumetric classification by processing axial slices through a pre-trained ResNet [5] and aggregating the resulting features using 3D convolutional layers. Trained on over 36,000 CT volumes from the RAD-ChestCT dataset, it demonstrated the feasibility of multi-abnormality prediction from unfiltered whole-volume data. CT-ViT [3], introduced as part of the GenerateCT framework, leverages a Vision Transformer to encode 3D CT volumes into a compact latent representation, enabling both generation and classification tasks. More recently, CT-Scroll [8] proposed a global-local attention mechanism that explicitly models the scrolling behavior of radiologists, combining coarse whole-volume context with fine-grained slice-level attention. CT-SSG [9] takes a complementary approach by representing CT volumes as structured graphs, where axial slice triplets serve as nodes processed through spectral graph convolutions, enabling the model to capture inter-slice dependencies while remaining computationally efficient.

Although these approaches achieve promising results, they predominantly operate on axial slices, potentially limiting their ability to capture complementary

anatomical patterns observable in coronal and sagittal views. In contrast, our method leverages all three orthogonal planes through a dedicated multi-view architecture designed to improve feature diversity and robustness in thoracic pathology classification.

Foundation Models The release of CT-RATE [4], a large-scale dataset pairing 3D chest CT volumes with radiology reports, has recently enabled the emergence of foundation models for CT imaging. CT-CLIP [4] leverages contrastive language-image pre-training for zero-shot multi-abnormality detection, while CT-FM [7] employs visual contrastive learning across multiple anatomical regions. Merlin [1] and COLIPRI [10] further combine vision-language alignment with additional self-supervised objectives on abdominal and chest CT data, respectively. While these foundation models excel at learning transferable representations, their visual encoders operate within a standard Euclidean framework and do not explicitly model structured inter-slice relationships or multi-view complementarity.

While foundation models aim to learn generic and transferable representations across diverse tasks and datasets, their broad scope may come at the expense of task-specific precision and often requires substantial computational resources for large-scale pre-training and deployment. In contrast, our work focuses on supervised visual encoder design, proposing a multi-view fusion architecture that can serve as a useful backbone for direct classification and may be beneficial for downstream foundation model development.

3 Method

3.1 Preprocessing

We conduct our experiments on chest CT scans from the CT-RATE dataset [4], and follow a preprocessing pipeline adapted from CT-Net [2] to prepare the 3D volumes as input to MV-CT-Net. This pipeline standardizes the CT volumes and enhances relevant anatomical structures through a sequence of intensity and spatial transformations. We modify the original pipeline to improve the preservation of global anatomical structures in chest CT volumes, which is important for enabling consistent representations across the axial, coronal, and sagittal views used in our multi-view framework.

We start by resizing the original CT volumes to a fixed spatial resolution of $480 \times 480 \times 240$. This is where we differ: we reshape the volumes using linear interpolation instead of cropping and padding them, which preserves the anatomical integrity of the scans while ensuring a consistent input size for the model. We conduct a study comparing the two approaches, as presented in the ablation study.

After this step, we apply the same preprocessing pipeline as CT-Net, which includes the conversion to Hounsfield Units (HU) with the slope and intercept

values provided in the DICOM metadata. This is followed by a windowing operation to focus on the relevant intensity range for thoracic abnormalities, which is between -1000 and 200 HU. Finally, we normalize the pixel values to the range [0, 1] to facilitate neural network training. The results of the preprocessing steps are illustrated in Figure 1.

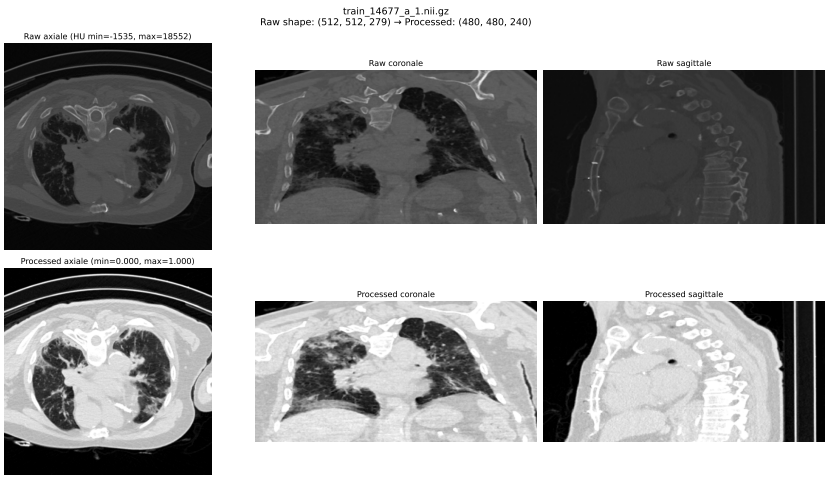

Fig. 1. Results of the preprocessing steps. The original CT volume is shown on the top row. The bottom row shows the result of the preprocessing steps.

3.2 Multi View Computed Tomography Network (MV-CT-Net)

The proposed Multi-View CT-Net (MV-CT-Net) processes 3D CT volumes $X \in \mathbb{R}^{D\times H\times W}$ by extracting features from three anatomical planes: axial, coronal, and sagittal. Each plane is processed through a dedicated branch of the network, which consists of a series of convolutional layers followed by a ResNet[5] backbone pre-trained on ImageNet. The resulting representations from each branch are then fused using a concatenation operation prior to classification. The design is motivated by the observation that different anatomical planes can provide complementary information, enhancing the model's ability to capture complex spatial relationships within CT volumes. An overview of the architecture is illustrated in Figure 2.

Feature Extraction Following CT-Net [2], we extend the feature extraction process to include the coronal and sagittal planes. For each anatomical axis, we group contiguous slices into triplets, forming 3-channel inputs compatible with a pre-trained 2D ResNet [5] backbone. Formally, given n slices along a given axis, we construct $\lfloor n/3 \rfloor$ triplet-based inputs.

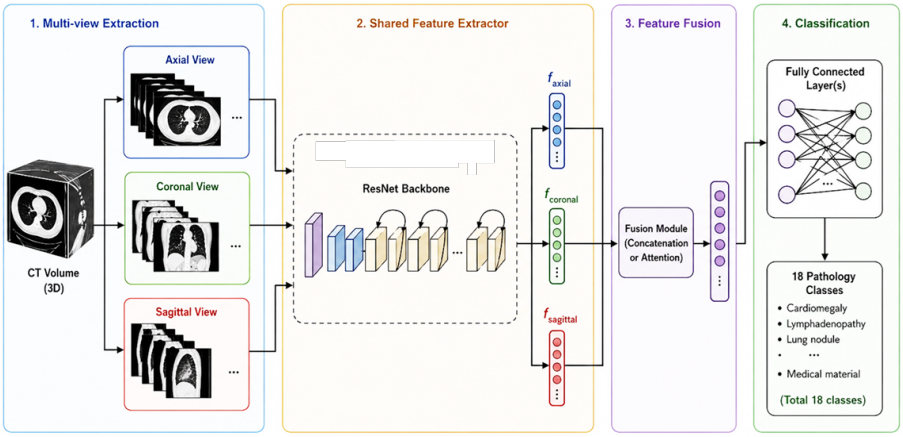

Fig. 2. Architecture of MV-CT-Net. The model processes CT volumes through three branches corresponding to the axial, coronal, and sagittal planes. Each branch consists of convolutional layers followed by a ResNet backbone, with the resulting features fused for classification.

For the axial plane, the input consists of 240 slices with spatial resolution 480×480, yielding 80 triplets. For the coronal and sagittal planes, we use 480 slices with spatial extents $D \times W$ and $D \times H$, respectively, resulting in 160 triplets per axis. Each triplet is encoded into a feature map of size $512 \times h_f \times w_f$, where $h_f = 15$ and $w_f = 15$ for the axial branch, and $h_f = 8$, $w_f = 15$ for the coronal and sagittal branches. The resulting features are stacked along the slice dimension, forming a 5D tensor of shape $(B, 512, S, h_f, w_f)$, where B denotes the batch size and S the number of triplets. Each axis-specific feature volume is then processed by a dedicated stack of 3D convolutional layers that progressively reduce spatial dimensions while compressing the channel dimension from 512 to 16. Finally, the resulting representations are flattened into 1D feature vectors. This yields a feature dimension of $d_{\text{axi}} = 7840$ for the axial branch, and $d_{\text{cor}} = d_{\text{sag}} = 8960$ for the coronal and sagittal branches.

Fusion and Classification The three per-axis feature vectors are concatenated into a single representation $z = [d_{\text{axi}}; d_{\text{cor}}; d_{\text{sag}}] \in \mathbb{R}^{d_{\text{total}}}$ where $d_{total} = d_{axi} + d_{cor} + d_{sag} = 25760$. This fused representation is then passed through a Multi-Layer Perceptron (MLP) consisting of two fully connected layers with ReLU activations and dropout ($p = 0.3$), followed by a final sigmoid output layer producing multi-label predictions for the 18 thoracic abnormalities..

Thresholds Learning To tackle the class imbalance in the CT-RATE dataset, we implement a threshold learning strategy where we optimize a set of class-specific thresholds on the training set to maximize the F1 score for each label. This approach is based on the work of CT-SSG[9] and allows us to convert the

predicted probabilities into binary classifications in a way that is tailored to the imbalanced class distributions. At the end of each training epoch, we take the sigmoid output probabilities in $[0, 1]$ for each of the n_{out} labels. Rather than using a fixed 0.5 threshold for all classes, we optimize a vector of thresholds $t \in [0, 1]^{n_{out}}$ by maximizing the F1 score on the training set. We then fix the optimized thresholds and apply them to the validation set to compute the binary predictions and evaluate the performance of the model using metrics such as AUROC, F1 score, precision and recall. Notice, we use the validation set as no separate test set is available in the CT-RATE dataset.

4 Experiments

4.1 Dataset

All models are trained and evaluated on the CT-RATE dataset[4], which consists of 50,188 reconstructed non-contrast 3D chest CT volumes from 21,304 unique patients. Each volume is annotated with 18 binary labels corresponding to the presence of specific thoracic abnormalities. These labels are automatically extracted from radiology reports using a pre-trained language model[11] that has been fine-tuned for this task. The dataset is split into 20,000 patients for training and 1,304 patients for the validation, following the official split provided with the dataset. No separate test set is available, and all experiments are conducted on this validation split. A comprehensive analysis of the dataset, shown in Figure 3, reveals significant class imbalance, with certain abnormalities being substantially more prevalent than others.

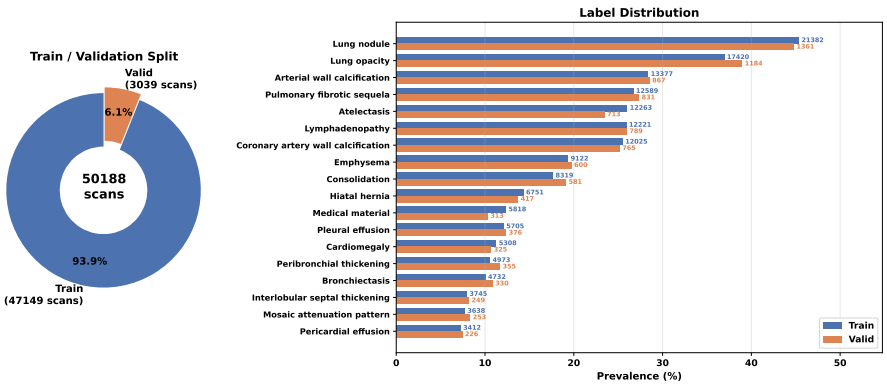

Fig. 3. Distribution of the 18 binary labels in the CT-RATE dataset. The dataset exhibits significant class imbalance, with some abnormalities being much more common than others.

4.2 Results and evaluation

We compare the proposed model with several baseline and state-of-the-art methods for multi-label chest CT classification, including CT-Net [2], CT-ViT [3], CT-Scroll [8], and CT-SSG [9]. We use the micro F1-score as the primary evaluation metric since it balances precision and recall, which is important for medical diagnosis tasks and better represents the label distribution. We also report AUROC, accuracy, precision, and weighted F1-score for a more comprehensive evaluation. The results are summarized in Table 1.

Method	AUROC	Accuracy	F1 score	W.F1	Precision
CT-Net[2]	82.40	**84.97**	48.34	42.67	**71.51**
CT-ViT[3]	73.92	70.83	45.01	49.65	35.59
CT-Scroll[8]	81.80	79.49	53.97	58.08	48.34
CT-SSG[9]	**83.64**	81.03	**59.19**	**60.27**	50.44
MV-CT-Net (ours)	<u>82.81</u>	<u>81.12</u>	<u>58.18</u>	<u>59.19</u>	<u>50.73</u>

Table 1. Evaluation on the CT-RATE validation set. We re-trained the CT-Net model on CT-RATE dataset. The best results are in bold, the second best are underlined.

We re-trained the CT-Net [2] model as a baseline. The original model was trained and evaluated on RAD-ChestCT; however, to ensure a fair comparison, we re-trained it on the CT-RATE dataset. We observe that MV-CT-Net achieves competitive performance compared with state-of-the-art methods. Notably, our model outperforms CT-Net on the CT-RATE dataset, achieving an improved F1 score

Although the CT-Net model achieves better performance in terms of accuracy and precision, we argue that this is largely due to class imbalance. The model tends to predict a majority of negative labels, which does not accurately reflect the true label distribution. This behavior is illustrated by the large discrepancy between the precision and F1 score of the CT-Net model.

For more detailed inspection, the per-label F1 scores for MV-CT-Net across all 18 abnormalities are reported in Figure 4.

We give a per label comparison of our model MV-CT-Net against CT-SSG in Figure 2. Notably, MV-CT-Net achieves the highest F1 score on 8 of 18 pathologies, with pronounced advantages on conditions characterized by diffuse parenchymal involvement where full volumetric multi-view reasoning is advantageous, such as pleural effusion, lung opacity, and emphysema. CT-SSG shows stronger performance on focal mediastinal findings such as pericardial effusion and hiatal hernia, where its explicit inter-slice relational modeling may better capture localized structural changes. These complementary strengths suggest that multi-view 3D aggregation and graph-based slice reasoning encode partially orthogonal information, opening avenues for hybrid approaches in future work.

Label	%Pos	MV-CT-Net	CT-SSG
Medical material	10.3%	0.5599	**0.6310**
Arterial wall calcification	28.5%	**0.7781**	0.7780
Cardiomegaly	10.7%	0.5949	**0.6240**
Pericardial effusion	7.4%	0.4019	**0.5540**
Coronary artery wall calcification	25.2%	0.7632	**0.7670**
Hiatal hernia	13.7%	0.3728	**0.4210**
Lymphadenopathy	26.0%	0.5107	**0.5200**
Emphysema	19.7%	**0.5232**	0.5090
Atelectasis	23.5%	0.4868	**0.5210**
Lung nodule	44.8%	**0.6421**	0.6370
Lung opacity	39.0%	**0.7401**	0.7380
Pulmonary fibrotic sequela	27.3%	**0.4913**	0.4670
Pleural effusion	12.4%	**0.8302**	0.8250
Mosaic attenuation pattern	8.3%	0.3780	**0.4560**
Peribronchial thickening	11.7%	**0.3992**	0.3940
Consolidation	19.1%	0.6070	**0.6410**
Bronchiectasis	10.9%	**0.4092**	0.3910
Interlobular septal thickening	8.2%	0.4118	**0.4180**
Micro Avg	–	0.5818	**0.5919**

Table 2. F1 Score Comparison: MV-CT-Net vs CT-SSG on CT-RATE validation set.

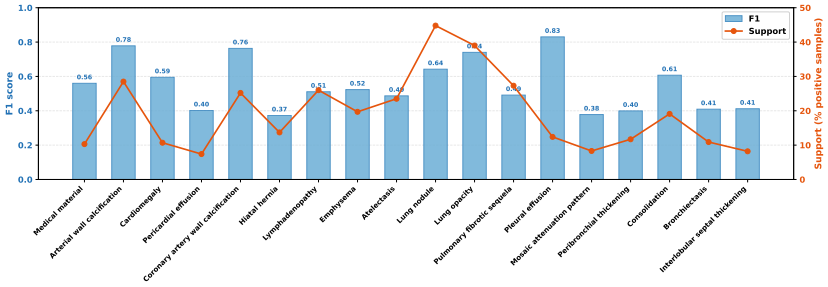

Fig. 4. CT-RATE validation performance (F1) across 18 labels with corresponding positive training samples.

4.3 Ablation study

In this section, we conduct an ablation study to evaluate the contribution of the different components of MV-CT-Net, as well as the impact of the proposed preprocessing and threshold learning strategies on performance.

First, the CT-Net [2] study highlighted that several target labels are mainly located in the liver region. However, with the original end-to-end preprocessing pipeline, this region may be partially cropped, leading to missed detections. To overcome this limitation, we propose a preprocessing strategy based on linear interpolation, in which all CT volumes are resized to a fixed resolution of $(240 \times 480 \times 480)$.

Method	AUROC	Accuracy	F1 score	Precision	Recall
Crop and pad[2]	**64.96**	80.12	57.02	48.85	68.46
Linear interpolation (ours)	62.61 ↓	**80.28**	**57.58** ↑	**49.15** ↑	**69.49** ↑

Table 3. Impact of preprocessing strategies on the performance. We compare our linear interpolation with the crop-and-pad method from CT-Net [2]. The model used in both methods is CT-Net for a quicker comparison.

To assess the effectiveness of this strategy, we compare it with the original crop-and-pad preprocessing used in CT-Net. Table 3 presents the impact of different preprocessing strategies on the performance of CT-Net. Compared with the crop-and-pad approach used in CT-Net [2], the proposed linear interpolation strategy provides slight improvements in overall classification performance, with the F1-score increasing from 57.02% to 57.58%, precision from 48.85% to 49.15%, and recall from 68.46% to 69.49%. A marginal improvement in accuracy is also observed, increasing from 80.12% to 80.28%. Although the AUROC decreases from 64.96% to 62.61%, the improvements in F1-score and recall suggest that the proposed interpolation strategy may help preserve more relevant anatomical information, allowing the model to better identify positive cases. These results indicate that maintaining a more complete volumetric context could be beneficial for improving model sensitivity in medical imaging applications.

Method	Accuracy	F1 score	Precision	Recall
No thresholds	**84.97**	48.34	**71.51**	36.52
Thresholds learning (ours)	80.28	**57.58**	49.15	**69.49**

Table 4. Impact of threshold learning on the performance. The model used in both methods is CT-Net for a quicker comparison.

Table 4 shows the impact of the proposed threshold learning strategy on the performance of CT-Net. Compared with the baseline model without threshold optimization, the proposed method improves the F1-score from 48.34% to 57.58% and significantly increases the recall from 36.52% to 69.49%. These results indicate that learning class-specific decision thresholds helps the model better identify positive samples, which is important in medical imaging applications where reducing false negatives is critical. However, the overall accuracy slightly decreases from 84.97% to 80.28%, while precision drops from 71.51% to 49.15%. This trade-off between precision and recall suggests that the proposed threshold learning strategy still requires further improvement to enhance the diagnostic sensitivity of the model.

5 Discussion

While MV-CT-Net achieves competitive results in terms of F1 score, accuracy, and precision, its AUROC (65.17) remains substantially lower than methods such as CT-SSG (83.64) and CT-Scroll (81.80). This gap suggests that the model's ranking of positive and negative predictions across all operating points can be improved, even though its binary decisions at the learned thresholds are effective. Investigating this discrepancy, for instance through calibration techniques or loss functions that directly optimize ranking performance, is an important direction for future work.

From a practical standpoint, the three-branch design triples the feature extraction cost relative to single-view methods, which may limit deployment in time-critical clinical workflows. Lightweight backbones or parameter sharing across branches could help mitigate this overhead. Additionally, the current evaluation is limited to a single dataset split of CT-RATE; cross-dataset generalization, as explored by CT-SSG [9], remains to be assessed.

A promising direction for future work is to leverage MV-CT-Net as a visual backbone within a contrastive learning framework, pairing multi-view CT representations with radiology reports in a vision-language pre-training paradigm similar to CT-CLIP [4]. The richer spatial encoding provided by multi-view fusion could yield stronger transferable representations for downstream tasks such as zero-shot abnormality detection and automated report generation.

6 Conclusion

We presented MV-CT-Net, a multi-view architecture for multi-label classification of 3D chest CT volumes that jointly processes axial, coronal, and sagittal planes through dedicated encoder branches. Combined with a linear interpolation preprocessing strategy and a class-specific threshold learning mechanism, MV-CT-Net achieves the highest F1 score (58.18), accuracy (81.12), and precision (50.73) on the CT-RATE benchmark, outperforming existing methods on the metrics most relevant to imbalanced clinical settings.

Future work will focus on addressing data imbalance through advanced sampling strategies, synthetic data generation, and cost-sensitive learning approaches. We also plan to investigate more effective fusion mechanisms between anatomical views and evaluate the proposed framework on additional large-scale medical imaging datasets to further assess its robustness and generalization capabilities. Future work may also explore attention-based fusion mechanisms across anatomical views and feature hierarchies to enable more adaptive integration of multi-view representations.

Acknowledgements The first author is financed by a CIFRE PhD granted by ANRT (2024/0992). We gratefully acknowledge CNRS and IDRIS for providing access to the Jean Zay supercomputer (AD010816084R1) for conducting the deep learning training.

References

1. Blankemeier, L., Cohen, J.P., Kumar, A., Van Veen, D., Gardezi, S.J.S., Parekh, M., Sber, S., Lungren, M.P., Langlotz, C.P., Chaudhari, A.S.: Merlin: A vision language foundation model for 3d computed tomography. arXiv preprint arXiv:2406.06512 (2024)
2. Draelos, R.L., Dov, D., Mazurowski, M.A., Lo, J.Y., Henao, R., Rubin, G.D., Carin, L.: Machine-learning-based multiple abnormality prediction with large-scale chest computed tomography volumes. Medical Image Analysis **67**, 101857 (Jan 2021). https://doi.org/10.1016/j.media.2020.101857
3. Hamamci, I.E., Er, S., Sekuboyina, A., Simsar, E., Tezcan, A., Simsek, A.G., Esirgun, S.N., Almas, F., Doğan, I., Dasdelen, M.F., et al.: Generatect: Text-conditional generation of 3d chest ct volumes. In: European Conference on Computer Vision. pp. 126–143. Springer (2024)
4. Hamamci, I.E., Er, S., Wang, C., Almas, F., Simsek, A.G., Esirgun, S.N., Dogan, I., Durugol, O.F., Hou, B., Shit, S., Dai, W., Xu, M., Reynaud, H., Dasdelen, M.F., Wittmann, B., Amiranashvili, T., Simsar, E., Simsar, M., Erdemir, E.B., Alanbay, A., Sekuboyina, A., Lafci, B., Kaplan, A., Lu, Z., Polacin, M., Kainz, B., Bluethgen, C., Batmanghelich, K., Ozdemir, M.K., Menze, B.: Generalist foundation models from a multimodal dataset for 3d computed tomography. Nature Biomedical Engineering (Feb 2026). https://doi.org/10.1038/s41551-025-01599-y, http://dx.doi.org/10.1038/s41551-025-01599-y
5. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (June 2016)
6. Li, X., Thrall, J.H., Digumarthy, S.R., Kalra, M.K., Pandharipande, P.V., Zhang, B., Nitiwarangkul, C., Singh, R., Khera, R.D., Li, Q.: Deep learning-enabled system for rapid pneumothorax screening on chest ct. European Journal of Radiology **120**, 108692 (2019). https://doi.org/https://doi.org/10.1016/j.ejrad.2019.108692, https://www.sciencedirect.com/science/article/pii/S0720048X19303420
7. Pai, S., Bontempi, D., Gidwani, M., Vaidya, A., Bressem, K., Perez-Johnston, R., Bou Zerdan, M., Perez-Lopez, R., Aerts, H.J.: A foundation model for clinical-grade computational pathology and rare cancer detection. Nature Medicine (2025), cT-FM: Vision foundation model for 3D CT
8. Piazza, T.D., Lazarus, C., Nempont, O., Boussel, L.: Imitating radiological scrolling: A global-local attention model for 3d chest ct volumes multi-label anomaly classification (arXiv:2503.20652) (Mar 2025). https://doi.org/10.48550/arXiv.2503.20652, http://arxiv.org/abs/2503.20652, arXiv:2503.20652 [cs]
9. Piazza, T.D., Lazarus, C., Nempont, O., Boussel, L.: Structured spectral graph learning for multi-label abnormality classification in 3d chest ct scans (arXiv:2510.10779) (Oct 2025). https://doi.org/10.48550/arXiv.2510.10779, http://arxiv.org/abs/2510.10779, arXiv:2510.10779 [cs]
10. Wald, T., Hamamci, I.E., Gao, Y., Bond-Taylor, S., Sharma, H., Ilse, M., Lo, C., Melnichenko, O., Schwaighofer, A., Codella, N.C.F., Wetscherek, M.T., Maier-Hein, K.H., Korfiatis, P., Salvatelli, V., Alvarez-Valle, J., Pérez-García: Comprehensive language-image pre-training for 3d medical image understanding (2026), https://arxiv.org/abs/2510.15042
11. Yan, A., McAuley, J., Lu, X., Du, J., Chang, E.Y., Gentili, A., Hsu, C.N.: RadBERT: Adapting transformer-based language models to radiology. Radiol Artif Intell **4**(4), e210258 (Jun 2022)

12. Zhang, G., Yang, Z., Gong, L., Jiang, S., Wang, L.: Classification of benign and malignant lung nodules from ct images based on hybrid features. Physics in Medicine & Biology **64**(12), 125011 (2019)