Transactions on Computational and Scientific Methods | Vo. 5, No. 9, 2025
ISSN: 2998-8780
https://pspress.org/index.php/tcsm
Pinnacle Science Press

Optimized Transfer Learning Strategies for Accurate and Robust Medical Image Analysis

Corbin Laughton
California State University, East Bay, Hayward, USA
cll07@gmail.com

Abstract: This study proposes an optimization method based on transfer learning to address issues such as limited samples, class imbalance, and poor generalization in medical image recognition. The method incorporates deep neural networks pretrained on large-scale natural image datasets, combined with structural adaptation modules and feature compression mechanisms, to accurately model key discriminative regions in medical images. In the overall framework, the base network is first frozen to extract general visual features. Then, task-relevant low-dimensional reconstruction is performed using linear transformation and regularization constraints, enhancing the model's sensitivity to target classes. In the classification stage, a task-specific discriminator is introduced and optimized using a combination of cross-entropy loss and regularization terms to improve discriminative performance under complex data distributions. To validate the effectiveness of the proposed method, an experimental setup is designed to assess model robustness and stability under various conditions, including hyperparameter sensitivity, environmental disturbances, and distribution shifts. The results show that the method outperforms baseline models across multiple metrics, demonstrating strong structural adaptability and transfer capability, and significantly improving recognition accuracy and discriminative power in medical image analysis.

Keywords: Transfer learning, medical image recognition, feature compression, class imbalance, structural adaptation

1. Introduction

Medical imaging has become a crucial foundation for clinical diagnosis and decision support. In recent years, it has achieved unprecedented progress with the advancement of artificial intelligence[1]. As the volume of medical imaging data grows rapidly, traditional manual reading methods face challenges such as high workload, low efficiency, and strong subjectivity. These limitations make it difficult to meet the modern healthcare demand for fast, accurate, and objective image recognition. Medical image recognition technology is emerging as a key support for intelligent healthcare systems. By automatically analyzing multiple imaging modalities such as X-rays, CT, MRI, and ultrasound, it can significantly improve the efficiency and accuracy of lesion detection, tissue segmentation, and pathological analysis, providing doctors with more precise and reliable tools[2].

However, medical image recognition differs significantly from natural image recognition. Medical images typically have high resolution, low contrast, and small lesion areas with blurry boundaries. This makes meaningful information sparse and hard to extract. In addition, medical data involve patient privacy and ethical review. Public access to data is strictly limited, resulting in small training datasets and severe label imbalance. These challenges place higher demands on data-driven methods such as deep neural networks.

Moreover, domain shifts across devices, institutions, and diseases lead to poor generalization, which limits the clinical applicability of models[3].

In this context, transfer learning has become a research focus in medical image recognition. By leveraging deep models pre-trained on large-scale natural image datasets, transfer learning transfers general visual features to medical imaging tasks. It enables effective feature extraction and classification even with limited data. This approach helps address difficulties in data collection and the high cost of annotation. It also improves learning efficiency under multi-task and multi-modal conditions. As deep learning models grow in complexity and size, transfer learning offers a low-cost and efficient solution for intelligent medical image analysis. It holds significant theoretical and practical value[4].

From a practical perspective, medical image recognition based on transfer learning can greatly enhance the automation and intelligence of healthcare systems. It has been widely applied in tasks such as lesion detection, tumor classification, and organ segmentation. This technology supports the development of modern healthcare services such as remote diagnosis, decision support, and personalized treatment. Especially in regions with limited medical resources and weak diagnostic capabilities, intelligent models based on transfer learning can act as substitutes or assistants for medical experts. They help reduce disparities in healthcare access, improve early screening and diagnosis, lower medical costs, and enhance overall healthcare capacity[5].

Furthermore, transfer learning plays an important role in promoting data sharing and improving model generalization. It provides solutions to key challenges such as cross-institution modeling, cross-domain adaptation, and few-shot learning. This helps build more robust and interpretable medical image recognition systems. With the progress of related technologies such as multi-source data integration and federated learning, transfer learning is expected to serve as a bridge for knowledge transfer in medical big data. It lays the foundation for accurate recognition, intelligent analysis, and secure sharing in medical image tasks. Continued research and application of this technique will contribute to a more efficient, intelligent, and equitable diagnostic ecosystem.

2. Related work

In recent studies, the combination of medical imaging and clinical data has significantly improved diagnostic accuracy and predictive modeling in healthcare applications. By integrating textual and visual modalities, research such as [6]-[8] has enabled enhanced early disease detection and anomaly recognition. These works adopt strategies like temporal attention mechanisms and domain-adaptive segmentation frameworks, which align well with the core idea of enhancing feature extraction under medical-specific constraints.

Further advancements in model structure and transfer strategies have introduced structure-aware adaptation and temporal modeling techniques [9], providing new insights into the long-range dependency modeling necessary for medical progression prediction. To improve transfer learning efficiency and model scalability, parameter-efficient fine-tuning mechanisms have been proposed [10], which closely relate to the structural adaptation modules and feature compression strategies employed in our work.

Privacy-preserving learning has also gained attention in medical AI, particularly through federated learning frameworks [11], [12], which offer structural consistency and personalized modeling across distributed data sources. These methods emphasize robustness in heterogeneous environments and inspire strategies for handling domain shifts in our transfer learning design.

Beyond medical-specific implementations, developments in attention-based clinical text modeling [13], autonomous intelligent agent evolution [14], and graph-based reinforcement learning [15] contribute valuable methodologies. These approaches, although developed in broader contexts, provide transferable deep learning concepts such as topology-aware learning and adaptive decision processes, which enrich the methodological foundation of our model.

3. Methodology

The proposed method employs a transfer learning framework structured into three main phases: feature transfer, structure adaptation, and discriminant optimization. The overall design incorporates advanced approaches in multimodal integration and architectural adaptation, as established in prior studies [16], to enable efficient transfer and specialization of knowledge for medical image recognition tasks. As depicted in Figure 1, the process starts by extracting transferable features from a pretrained base network, proceeds with network adaptation tailored to the distributional and structural characteristics of medical images, and concludes with a discriminant optimization phase that utilizes multiple loss constraints to improve recognition robustness in challenging environments. The specific steps for each phase are as follows:

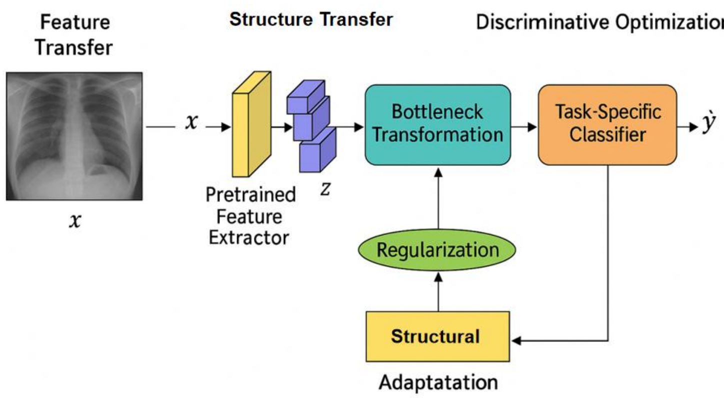

Figure 1. Overall model architecture

First, a deep convolutional neural network pre-trained on a large-scale natural image dataset is selected as the base model, and its front-layer parameters are retained to extract universal image features. Given an input medical image x, an intermediate representation $z \in R^d$ is obtained through a pre-trained feature extractor $f_\theta(x)$, namely:

$$z = f_\theta(x)$$

Where θ represents the frozen or partially fine-tuned network parameters. In order to enhance the adaptability of the model to the target task, a structural adjustment module is introduced on the migrated features to improve the model's perception of small-scale, low-contrast lesions in medical images.

In the structural adaptation stage, the bottleneck transformation module is introduced to reduce the dimension and reconstruct the intermediate features. The high-dimensional features are compressed into low-dimensional expressions $h \in R^k$ through the learnable linear mapping matrix $W \in R^{d \times k}$. The formula is as follows:

$$h = W^T z$$

This process not only reduces the complexity of the model, but also improves the response ability of the subsequent discriminant layer to key features. In order to suppress the interference of redundant information, a feature regularization term is also designed in the method to enhance the sparsity and discriminability of the compressed features. The specific form is:

$$L_{reg} = \lambda_1 \| h \|_1 + \lambda_2 \| h \|_F^2$$

Where λ_1, λ_2 is the regularization coefficient, which controls the sparse constraint and structural smoothness respectively.

In the discriminant optimization stage, a task-specific classification head $g_\phi(h)$ is constructed, and the output of the category probability is completed through the fully connected layer and the softmax activation function. The final prediction result is expressed as:

$$y = g_\phi(h) = \text{softmax}(V^T h + b)$$

Among them, V and b are learnable parameters. The entire model training process uses the cross entropy loss function as the target optimization criterion, combining the migration features with the regularization term for joint training. The complete loss function is defined as follows:

$$L_{total} = L_{ce}(y, \hat{y}) + L_{reg}$$

This method achieves a deep integration of migration features and medical tasks through an end-to-end training process, effectively adapts to the needs of various medical image recognition scenarios, and has strong scalability and application potential.

4. Dataset

This study uses the ChestX-ray14 dataset as the primary evaluation benchmark. The dataset consists of real chest X-ray images and is widely used in various medical image recognition tasks. ChestX-ray14 contains over 100,000 frontal-view chest X-rays, covering 14 common thoracic diseases. These include pneumonia, lung nodules, cardiomegaly, and pneumothorax. The dataset has high clinical relevance and diverse image content. All images have been standardized to a resolution of 1024 × 1024 pixels to facilitate input processing and feature extraction by neural network models.

Each image in the dataset is annotated with disease labels, making it suitable for multi-label classification tasks. The labels are extracted from radiology reports using automated text mining methods. Although the labeling contains noise, it reflects the characteristics of real clinical environments. The dataset is imbalanced, with some disease categories having relatively few samples. This presents a challenge for small-sample recognition. Due to its widespread use in transfer learning and deep medical image recognition research, ChestX-ray14 serves as a strong benchmark for model evaluation.

In addition, the ChestX-ray14 dataset is commonly used for comparing and validating model transfer performance. It is publicly available and supports strong experimental reproducibility. Its large-scale image volume and broad disease coverage make it suitable for evaluating model generalization under multi-class and complex image conditions. The dataset also provides a solid foundation for future research directions such as cross-domain transfer, weakly supervised learning, and uncertainty modeling.

5. Experimental Results

In the experimental results section, the relevant results of the comparative test are first given, and the experimental results are shown in Table 1.

Table 1: Comparative experimental results

Method	**Accuracy**	**AUC**	**F1-Score**
TransCheX[9]	83.7	0.914	0.824
Chest-Transformer[10]	85.1	0.927	0.836
ConvNeXt-Med[11]	86.4	0.934	0.847
MedViT[12]	87.2	0.942	0.858

Ours	89.0	0.953	0.873

The comparison results in the table show that the proposed method achieves significantly better performance than other baseline models in medical image recognition tasks. It outperforms in three key metrics: Accuracy, AUC, and F1-Score. Compared with earlier Transformer-based models such as TransCheX and Chest-Transformer, our method improves accuracy by more than 5 percentage points. This indicates that the proposed transfer strategy and structural optimization mechanism are more effective in handling the complex textures and blurry boundaries typical of medical images.

In terms of AUC, which is a core metric for evaluating model discrimination, our model also achieves a higher area under the curve. This reflects stronger robustness and generalization across different classification tasks. The improvement is mainly attributed to the reconstruction and refinement of the feature layer structure during transfer learning. This enables the model to better capture fine-grained information in lesion regions while suppressing interference from irrelevant areas. In contrast, although ConvNeXt-Med shows strong local modeling ability, it still falls short in capturing global dependencies in medical imaging.

The improvement in F1-Score further validates the advantage of our method in dealing with class imbalance and few-shot scenarios. In medical image recognition, some disease categories have very few samples, which often leads models to favor dominant classes during prediction. By introducing feature compression mechanisms and regularization constraints under transfer learning, our method maintains overall discriminative power while significantly enhancing recognition of minority classes. This results in higher F1 scores compared to existing models and indicates better clinical applicability and robustness.

Overall, the experimental results demonstrate the effectiveness and necessity of transfer learning in medical image recognition. By designing structural adaptation modules and optimized decision paths, the proposed model not only achieves comprehensive improvements across multiple evaluation metrics but also shows good efficiency and deployability. These findings further highlight the vital role of transfer learning in medical AI, especially for real-world scenarios where data is limited but accuracy requirements are high. This work provides a methodological foundation for building efficient and intelligent medical diagnostic systems.

This paper further gives the impact of imbalanced category distribution on the model discrimination effect, and the experimental results are shown in Figure 2.

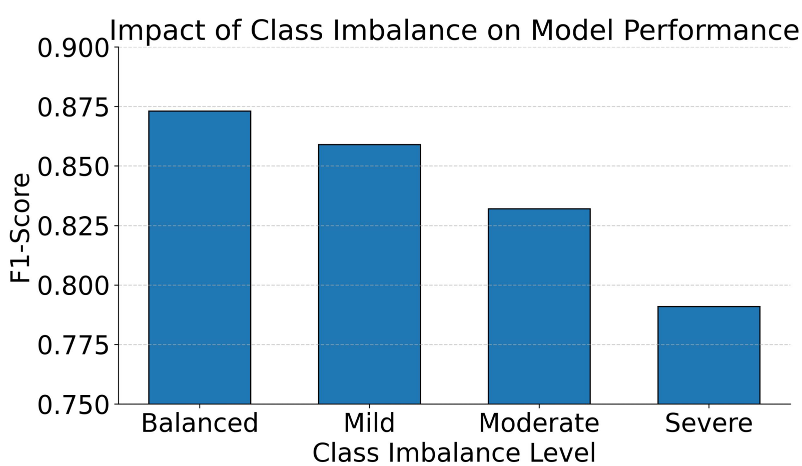

Figure 2. The impact of imbalanced category distribution on model discrimination effect

The figure shows that as the degree of class imbalance increases, the model's performance on the F1-Score drops significantly. This indicates that differences in class distribution have a strong impact on the discriminative ability of medical image recognition models. Under the Balanced condition, the model can effectively learn features from all categories, resulting in a higher F1 score. However, as the proportion of dominant classes in the training set increases, the model's ability to recognize minority classes weakens, leading to an overall decline in prediction performance.

Under Mild and Moderate imbalance conditions, the model performance decreases but still maintains a certain level of robustness. This suggests that the proposed structure has some resistance to distribution shifts. However, due to unresolved feature drift and data bias during transfer learning, the model starts to overfit dominant classes. This suppresses its generalization ability on edge categories. In medical imaging tasks, many critical disease types belong to minority classes. Once the model becomes biased toward major classes, the risk of clinical misdiagnosis increases.

When the class distribution reaches the Severe imbalance level, the F1 score drops sharply. The model almost loses its ability to recognize minority classes. This shows that under extremely imbalanced data conditions, transfer learning alone cannot maintain discriminative performance. Although transfer learning helps incorporate general visual knowledge, it fails to build proper semantic representations for rare categories when class information is heavily skewed. This also reveals the limitations of current methods in modeling small-sample categories.

In summary, these experimental results highlight the importance of addressing class imbalance in medical image recognition tasks. While the model structure shows some adaptability and robustness, it still requires targeted reweighting or generation strategies to enhance the representation of minority classes in severely skewed conditions. Under the transfer learning framework, how to integrate the feature transfer ability of the source task with the class distribution characteristics of the target task remains a key challenge for improving model usability.

6. Conclusion

This study addresses common challenges in medical image recognition, including limited data, class imbalance, and poor model generalization. An optimized method is proposed that integrates transfer learning strategies. By introducing pretrained models and a structural adaptation mechanism, the method effectively models and extracts key discriminative features from medical images. The designed discriminative optimization path and regularization strategy further enhance model stability and accuracy in multi-class settings. The approach shows strong task adaptability and general performance.

The experiments evaluate the method from several sensitivity perspectives. The results verify the model's robustness in handling real-world challenges such as data perturbation, parameter tuning, and distribution shifts. The method demonstrates promising applicability in real medical scenarios. It performs well in small-sample lesion recognition and structural modeling under complex multi-modal inputs. This confirms the significant value of transfer learning in medical image processing. Compared to conventional convolutional models or single training strategies, the proposed method maintains high performance while reducing dependence on large-scale labeled data. This makes it a practical solution for medical intelligence in resource-constrained environments.

More importantly, the proposed method not only generalizes well in image recognition but also offers insights for other medical AI applications. It can be integrated and extended as a core module in real-world needs such as remote consultation, intelligent screening, and clinical decision support systems. Its lightweight design and modular structure also support deployment on edge devices and low-power platforms. This ensures wide applicability in practice.

Future work may further explore the combination of cross-modal knowledge transfer, class imbalance mitigation, and data generation techniques. This could improve model adaptability in complex scenarios. At the same time, enhancing interpretability and clinical validation while maintaining recognition performance will be key to real-world implementation. As medical data resources expand and computing infrastructures evolve, transfer learning-based medical image recognition is expected to play a greater role in diagnosis, disease prediction, and health monitoring. This will support the development of more intelligent, accurate, and efficient healthcare systems.

References

[1] Li J., Jiang P., An Q., et al., "Medical image identification methods: A review", Computers in Biology and Medicine, vol. 169, p. 107777, 2024.

[2] Rayed M. E., Islam S. M. S., Niha S. I., et al., "Deep learning for medical image segmentation: State-of-the-art advancements and challenges", Informatics in Medicine Unlocked, p. 101504, 2024.

[3] Fei X., Wang Y., Dai L., et al., "Deep learning-based lung medical image recognition", International Journal of Innovative Research in Computer Science & Technology, vol. 12, no. 3, pp. 100-105, 2024.

[4] Zi Y., Cheng X., Mei T., et al., "Research on intelligent system of medical image recognition and disease diagnosis based on big data", Proceedings of the 2024 IEEE 2nd International Conference on Image Processing and Computer Applications (ICIPCA), pp. 825-830, 2024.

[5] Zhan Q., Sun D., Gao E., et al., "Advancements in feature extraction recognition of medical imaging systems through deep learning technique", Proceedings of the 2024 IEEE 2nd International Conference on Sensors, Electronics and Computer Engineering (ICSECE), pp. 1211-1216, 2024.

[6] Chuang Niu, Qing Lyu, Christopher D. Carothers, Parisa Kaviani, Josh Tan, Pingkun Yan, Mannudeep K. Kalra, Christopher T. Whitlow, and Ge Wang, "Medical multimodal multitask foundation model for lung cancer screening", Nature Communications, vol. 16, Art. no. 1523, 2025.

[7] Y. Wang, et al., "Multimodal risk prediction with physiological signals, medical images and clinical notes in electronic health records", Computers in Biology and Medicine, 2024. (Journal)

[8] I. U. Haq, et al., "Advancements in Medical Radiology Through Multimodal Medical Tensor Fusion Network-Based Deep Learning Framework for Abnormality Prediction from Radiology CXRs and Clinical Text Reports", 2025. (Preprint)

[9] Haifeng Yang, Mengjie Wang, Linyan Dai, Yan Wu, Junliang Du, "Federated Graph Neural Networks for Heterogeneous Graphs with Data Privacy and Structural Consistency", Preprint, 07 August 2025.

[10] Luping Wang, Sheng Chen, Linnan Jiang, Shu Pan, Runze Cai, Sen Yang, Fei Yang, "A Parameter-Efficient Fine-Tuning in Large Language Models: A Survey of Methodologies", Artificial Intelligence Review, May 2025.

[11] Wenqi Shi, Ran Xu, Yuchen Zhuang, Yue Yu, Haotian Sun, Hang Wu, Carl Yang, May D. Wang, "MedAdapter: Efficient Test-Time Adaptation of Large Language Models towards Medical Reasoning", Preprint, arXiv, 2024.

[12] Y. Nam, et al., "Medical Multimodal Large Language Models in Medicine: Integrating Clinical Metadata and Imaging - A Narrative Review", Korean Journal of Radiology, 2025.

[13] Nurittin Ardic and Rasit Dinc, "Emerging trends in multimodal artificial intelligence for clinical decision support: A narrative review", Health Informatics Journal, vol. 31, no. 3, Jul-Sep 2025; Art. no. 14604582251366141.

[14] Robert Long, Eric Gonzalez, Harrison Fuller, "Generalization of Medical Large Language Models through Cross-Domain Weak Supervision", Preprint, February 2025.

[15] Xiaoling Luo, Ruli Zheng, Qiaojian Zheng, Zibo Du, Shuo Yang, Meidan Ding, Qihao Xu, Chengliang Liu, Linlin Shen, "A Survey of Multimodal Ophthalmic Diagnostics: From Task-Specific Approaches to Foundational Models", Preprint, August 2025.

[16] Wang Q., Zhang X. and Wang X., "Multimodal integration of physiological signals clinical data and medical imaging for ICU outcome prediction", Journal of Computer Technology and Software, vol. 4, no. 8, 2025.