

Received: 09 Aug 2025, Accepted: 10 Sept 2025, Published: 12 Sept 2025

DOI: <https://doi.org/10.63503/j.ijcma.2025.161>

Review Article

Self-Supervised Learning Frameworks for MRI-Based Medical Image Classification: A Systematic Review

Tahseen Ali¹, Faiyaz Ahmad¹, Musheer Ahmad¹

¹ Department of Computer Engineering, Jamia Millia Islamia, New Delhi, India

Email: tahseenali834@gmail.com, fahmad1@jmi.ac.in, mahmad9@jmi.ac.in

*Corresponding Author: Tahseen Ali, tahseenali834@gmail.com

Abstract

In recent years, Self-supervised learning (SSL) has evolved as a framework that eliminates the bottleneck of dependence on labeled data for conventional Supervised learning paradigms. It has opened entirely new possibilities by removing the dependence on the annotated data, which is particularly useful for medical image classification problems, as finding the annotated data is scarce in this domain. In this paper, we have collected and presented a structured view of the usage of SSL framework and techniques for MRI-based medical image classification. We have consolidated a detailed review of this modality that can be practical for future research and experimentation.

Keywords: Machine Learning, Self-supervised Learning, Medical Image Analysis, Classification, MRI

1 Introduction

In modern diagnostic radiology, one of the most important aspects is MRI - Medical Resonance Imaging. MRIs provide a detailed structural and functional insight into the human body. In recent times where speed and scalability both matter, there has been a constant need for automated systems that can help radiologists classify and interpret these images. The rise in the demand is also because of the growing volume of MRI based images and need for their accurate diagnosis. The conventional Supervised techniques do not scale as they have a huge dependence on the annotated data. This greatly restricts the development of robust deep learning models for interpreting MRI [1]. In recent times, the Self-supervised learning framework has evolved as one of the promising frameworks, eliminating the dependence on the annotated data as it finds the cues within the data itself. Unlike traditional Supervised Learning, SSL employs natural structures or features of the data to generate supervision cues for the representation learning. This field, besides transforming Computer Vision and Natural Language Processing (NLP), is also being explored increasingly for medical image analysis, such as MRI [2, 3]. SSL works on the idea of using a smaller proxy task for representation learning. These proxy tasks are called Pretext tasks. The Pretext tasks are frameworks designed to train models to solve some problem and while solving the problem, it also learns the features of the data. These tasks have various types including distinguishing whether the two different rotations of an image are similar or dissimilar, reconstructing masked sections of an image, putting image patches back in the order, and so on. These tasks are categorized as Generative, Contrastive, etc. These trained models are then used for the actual downstream tasks like classification, segmentation, and so on [2]. This approach is particularly useful for Medical Image classification as 3D volumes contain rich spatial and temporal data that can be used to learn from with the appropriate Pretext tasks. Recent work in this domain has labeled SSL as one of the efficient frameworks for a variety of tasks like MRI-based classification, Schizophrenia classification, Alzheimer diagnosis, etc. These experiments illustrate how SSL can enhance transfer learning between tasks, generate more generalizable representations, and reduce dependency

on costly annotations [1, 3]. In this work, the most recent advancements in Self-supervised learning for MRI-based image classification are compiled, along with a description of key methods, a classification of Pretext problems, and trends in performance gains.

2 Self-Supervised Learning Approaches and Pretext tasks

Self-supervised Learning (SSL) has become an important approach in the wider field of Unsupervised learning. It aims to use the natural structure of data to learn meaningful representations without needing manual labels. The idea of SSL dates back to earlier studies like [4].

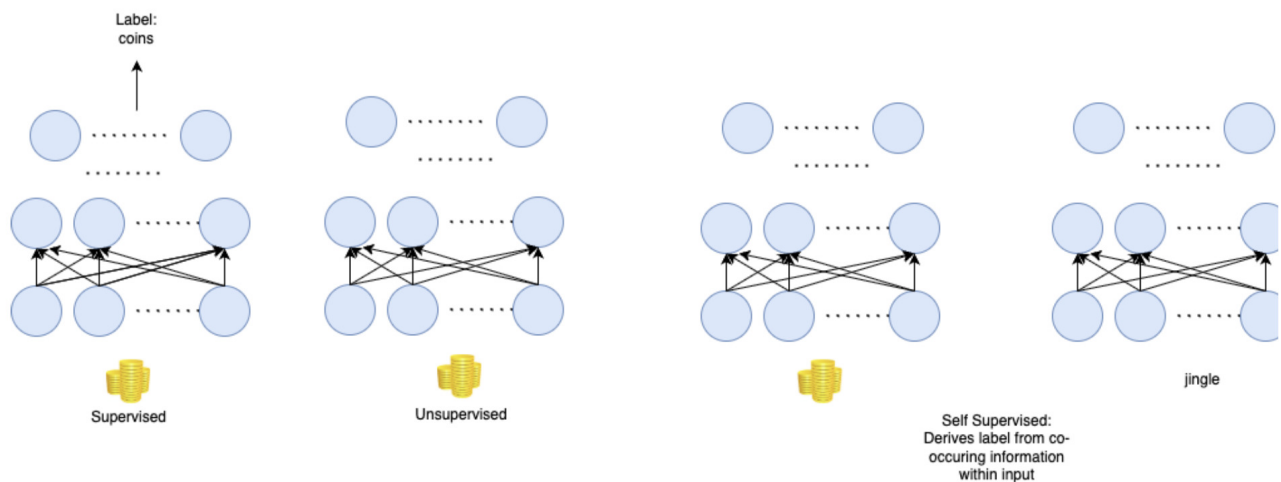


Figure 1: Self-supervised Learning Workflow

These studies introduced learning in real-world situations with different types of inputs. For instance, an image of a pile of coins might not always have a label, but it often is associated with the sound "jingle." This shows how valuable it is to learn from the relationships between different types of data. Over time, the machine learning community has made great progress in SSL, recognizing its ability to extract useful signals from the data itself. Unlike traditional Supervised learning, SSL creates output labels from within the data by revealing connections between different parts or views of the input data. For instance, Autoencoders (AEs) aim to reconstruct input data from compressed versions and can be seen as early types of SSL, since their output labels are simply the original data. AEs are commonly used for tasks like reducing dimensionality and detecting anomalies. SSL can be understood as: (1) predicting one part of the input from another, (2) predicting the future based on the past, (3) predicting what we cannot see from what we can see, and (4) reconstructing hidden, masked, or damaged parts from the visible content. At its heart, SSL involves hiding part of the input and training a model to predict or fill in the missing pieces. Jing et al. [5] expanded the definition of SSL to include any learning method that does not rely on human-made labels. With this definition, SSL becomes similar to Unsupervised learning. This includes models like Generative Adversarial Networks (GANs) [6], which learn to create realistic data through adversarial training without labeled supervision. A key innovation in SSL is using Pretext tasks, which are side tasks not linked to the model's end goal. These tasks help the model learn useful representations. They offer indirect guidance, helping the model find patterns in the data. Pretext task plays a crucial role in representation learning. They enable self-supervision framework to work by generating labels (cues) directly from the data. The significant benefit is that this eliminates the need for huge annotated data sets.

2.1 Pretext tasks and Representation Learning

2.1.1 Representation learning from the innate image features

In this method, the innate structure of images is leveraged, which is quite natural to the medical images. Some examples of this kind of Pretexts include solving a jigsaw puzzle, reordering mixed image patches, sequence remembering, etc. [1, 2].

2.1.2 Generative Pretext Learning

In this Pretext task, the models are trained to recreate the input images. This kind of tasks involve autoencoders (AEs), variational auto encoders (VAEs), or GANs (generative adversarial networks) to train models to create useful presentations. As they focus on image reconstruction, this allows them to perform classification without the need for manual annotations [2].

2.1.3 Contrastive Representation Learning

This is one of the most explored fields with very promising use cases. In this learning framework, the model is trained to tell apart the dissimilar images (negative pair) and pull together the similar images (positive pair). Negative pairs come from the dissimilar images whereas positive pairs come from the same image with different augmentations. Methods like SimCLR, MoCo, BYOL, and SimSiam use this approach to gain transformation-invariant, high-level features that work well across medical imaging tasks [7].

2.1.4 Predictive Masked Modeling

In this Pretext method, the model learns by predicting the masked part of the images. Some parts of the images are being covered and then the model is being trained to guess the missing content. This is the same as the method used in Natural Language Processing. This approach often uses vision transformers, such as MAE and BEiT, to learn detailed but context-aware representations [8]. The emphasis is on recovering parts of the image using visual information, rather than reconstructing the entire image like generative models do.

3 Literature Review

Magnetic Resonance imaging is a key imaging technique in medical image classification using SSL. MRI is a technique that has its use cases in fields like neurology, psychiatry, cardiology and musculoskeletal studies due to its non-invasive nature and high image quality of soft tissues. Since labelled data is costly and research direction is shifting to SSL to eliminate the need for labelled data, this framework has led to the utilisation of large amounts of unlabelled data and also improved model performance. In this section we have reviewed the recent work in a collection of 17 studies which have shown methodological improvements and clinical goals. This approach includes generative, contrastive and context-predictive learning, often combining elements of different methods to leverage the performance.

Several studies used contrastive learning, in which models learn to tell apart similar and different image representations. To learn features from the unlabeled MRI, Firildak et al. (2024) [9] proposed a valuable solution. He proposed using the SimCLR model to learn features from unlabeled MRI scans. Their training results were very impressive with 96 of accuracy and 72 of validation accuracy. The dip in the validation accuracy highlights the challenge of generalisation in SSL techniques. In 2021, Dufumier et al. (2021) [10] proposed in their research work a modified SimCLR approach to classify the psychiatric imaging tasks. They got an improvement in AUROC from 0.942 to 0.968 for schizophrenia and bipolar disorder classification. Similarly, Li et al. (2021) [11] in their research work, employed a contrastive method for brain tumor classification and obtained a boost in sensitivity from 0.888 to 0.920. In [12] Manna et al. (2021) demonstrated the significance of how even minor gains in performance (AUROC: 0.970 vs. 0.965) in the detection of anterior cruciate ligament (ACL) tears on knee MRIs were achievable using semi-supervised learning rather than fully supervised approaches. Similarly in [11], Hongwei Li employed contrastive learning on MRI data for detecting brain tumors in their work published in 2021.

An encoder extracted features, which were then classified using an SVM. This achieved a sensitivity of 0.92 and an AUC of 0.888, indicating strong diagnostic performance. Other research, like that by Fedorov et al. (2021) [13], looked at mutual information maximization for detecting Alzheimer's. Although the SSL-based model performed slightly worse (AUROC: 0.841) than the supervised model (0.880), it demonstrated the potential of SSL in important diagnostic tasks. Ouyang et al. (2021) [14] made progress by using a longitudinal neighborhood embedding structure with autoencoding. This approach improved Alzheimer's classification accuracy, increasing it from 0.794 to 0.836.

In the area of generative modeling, researchers focused on autoencoders and their ability to reconstruct image data to learn latent representations. Zhao et al. (2020) [15] introduced a generative autoencoder for predicting brain age and detecting Alzheimer's. In this method the researchers put in place a latent space constraint, which achieved a 16.6 improvement over supervised baselines. Similarly Osin et al. (2020) [16] in their work combined autoencoders and LSTMs. Their goal was to capture the temporal dynamics in functional MRI (fMRI) data related to psychiatric traits like PTSD. However they did not fully publish the quantitative metrics, but their work gave a new direction to temporal learning. In another work, Fedorov et al. (2021) [17] combined autoencoders with SimCLR to extract MRI features for Alzheimer's diagnosis. They achieved this through a linear classifier, however they didn't share the comparative result benchmarks. Manna et al. (2021) [12] created a jigsaw-based pretext task to classify ACL tears in knee MRIs and achieved an AUROC of 0.848. Hashimoto et al. (2021) [18] studied schizophrenia with a domain-specific transformation task, reporting an accuracy of 0.778. Dezaqi et al. (2021) [19] made progress by modeling temporal cycle consistency in MRIs derived from echocardiograms. This led to improved accuracy from 0.609 to 0.787. Zhu et al. (2021) [20] used simple transformations like rotation and patch ordering on COVID-19-related MRI and CT scans, but they did not provide separate details on MRI outcomes. Han et al. (2023) [21] addressed a common problem in self-supervised learning called representation collapse. This is when the model does not learn meaningful features. They introduced an Enhanced Masked Image Modeling (E-MIM) framework that combines various masking techniques with a contrastive regularization component. This method resulted in significant improvements in segmentation accuracy for multimodal MRI scans. More recent studies have started to combine different SSL strategies into unified frameworks. Haghighi et al. (2021) [22] created a combined approach that included generative modeling, patch restoration, and innate relationship tasks with pseudo-labeling to improve diagnostic accuracy on CT and MRI data. Their model achieved a strong AUROC improvement from 0.943 to 0.985.

Hsieh et al. (2020) [23] used a multimodal setup that combined autoencoders with contrastive objectives to investigate cognitive impairment and Alzheimer's, reaching a classification accuracy of 0.594; however, they did not provide baseline comparisons. A few broader patterns appear from these findings. Neurological disorders, especially Alzheimer's disease and Schizophrenia, dominate the clinical application space, followed by musculoskeletal and cardiovascular conditions. Among all approaches, Contrastive Learning was the most commonly used and generally led to modest but consistent performance improvements of around 6. When designed carefully, generative models provided the biggest performance improvements, achieving gains as high as 24.5 compared to Supervised methods. Intrinsic proxy tasks with clinical relevance, such as modeling heart motion or brain structure, often outperformed those based on artificial changes like patch shuffling. This highlights the benefit of adjusting learned representations directly to the downstream task. In conclusion, Self Supervised learning has shown great potential in medical image classification using MRI. While different SSL approaches have their own strengths, the overall evidence suggests that combining strategies, designing clinically relevant pretext tasks, and fully fine-tuning can significantly improve model performance in data-limited medical fields. Table 1 presents the summary of major works in a tabular form.

In medical imaging, especially for MRI-based classification, SSL tackles an important issue of the lack of labeled data. Labeling medical images is often expensive and requires expert input. SSL offers a promising solution by using large amounts of unlabeled scans to pre-train models. These models are then fine-tuned on a few annotated examples. This method is particularly useful in healthcare, where rules and ethical concerns often limit large-scale labeling efforts.

4 How Well Does SSL Work?

Most studies agree that SSL improves performance. Models pre-trained with SSL usually outperform those trained from scratch or with ImageNet weights, especially when there is limited labeled data. Reported improvements vary widely, reaching up to 32 in AUROC and nearly 30 in accuracy. This shows how much SSL can help in settings

Table 1: Summary of Selected Studies on Self-Supervised Learning for MRI

Author(s)	Year	Method / SSL Type	Dataset / Modality	Results / Metrics
Firildak et al. [9]	2024	Contrastive Learning (SimCLR)	MRI	Train acc: 96%, Val acc: 72%
Dufumier et al. [10]	2021	Modified SimCLR (Contrastive)	MRI	AUROC: 0.968
Li et al. [11]	2021	Contrastive	MRI	Sensitivity: 0.920
Manna et al. [12]	2021	Semi-supervised	MRI	AUROC: 0.970 vs. 0.965 (baseline)
Hongwei Li [11]	2021	Contrastive, SVM	MRI	Sensitivity: 0.92, AUC: 0.888
Fedorov et al. [13]	2021	Mutual information maximization	MRI	AUROC: 0.841 vs. 0.880 (supervised)
Ouyang et al. [14]	2021	Longitudinal neighborhood embedding, Autoencoder	MRI	Acc: 0.836
Zhao et al. [15]	2020	Generative Autoencoder	MRI	16.6% improvement over supervised methods
Manna et al. [24]	2022	Jigsaw pretext task	MRI	AUROC: 0.848
Hashimoto et al. [18]	2021	Domain-specific transformations	MRI	Accuracy: 0.778
Dezaki et al. [19]	2021	Temporal cycle consistency	MRI	Acc: 0.787
Han et al. [21]	2024	Enhanced Masked Image Modeling (E-MIM), Contrastive Learning	MRI multi-modal	Significant segmentation improvement
Haghighi et al. [22]	2021	Combined (generative, patch restoration and pseudo-labeling)	CT and MRI	AUROC: 0.985
Hsieh et al. [23]	2020	Multimodal AE, Contrastive Learning	MRI and multimodal	Accuracy: 0.594

with few resources. A few studies did report slight decreases in performance, often because of mismatched pretext tasks or weak fine-tuning setups. Still, the overall picture is clear: SSL reduces the need for extensive annotation and encourages broader use of machine learning in real clinical settings.

5 Why Does Medical Data Need Careful Pretext Design?

Many SSL methods used in medical imaging are based on Contrastive Learning, such as SimCLR, MoCo, and BYOL. These methods perform well with natural images, but medical images have unique challenges. In Contrastive Learning, small changes, like cropping or masking, are thought to preserve the image's meaning. However, in an MRI scan, cropping might remove a tumor or an important feature. This can change the diagnostic content completely. To address this, researchers are developing better augmentations. One example of better augmentation is Window Slicing and Window Warping in fMRI time series data [25]. In rs-fMRI (resting state analysis), the continuous time series data of brain activity is recorded by segmenting into overlapping temporal windows, which is called Window Slicing. Or it is being warped at the two end points (beginning and ending) which is called Window Warping. This

helps in maintaining temporal consistency and retaining important neurological information within the data [26]. In pathology, stain normalization and tile-based consistency are also important. These specific strategies show that semi-supervised learning in healthcare needs more than just applying concepts from natural image tasks.

6 Mixing SSL Approaches for Better Results

Several self-supervised goals are combined in some of the best models. For example, contrastive learning can be used along with simpler tasks like image jigsaw puzzles or rotation prediction. By teaching the model both the overall structure and specific details, these mixed approaches seem to improve performance. Although publication bias may have played a role in some of this success, the results are still promising. Using multi-task pretext techniques might offer additional benefits for complex modalities like MRI, which rely on spatial and contextual consistency.

7 How Fine-Tuning Affects Performance?

How you adjust SSL-pretrained models matters when using them. Research shows that optimizing the entire model, rather than just the last layer, often leads to better results. Tuning the whole network to the new data is beneficial because hospitals use different MRI protocols, scanners, and settings. This holds true even when labeled data is limited, suggesting that starting with a solid SSL foundation may reduce overfitting issues. Some researchers are exploring middle-ground strategies, like using different learning rates for each layer or unfreezing layers gradually.

8 Current and Emerging Applications of SSL

Up to now, the majority of SSL work has been radiology-related, particularly chest imaging. The availability of huge datasets like CheXpert and MIMIC-CXR is primarily to blame for this. However, results from chest X-rays might not be transferable to other fields. There are problems with MRI, CT, and ultrasound, particularly when dealing with 3D and 4D data. MRI varies widely in imaging sequences and resolution, making it hard to create general SSL models. Additionally, there are not many public MRI datasets available. SSL has been somewhat successful in fields like pathology and dermatology, particularly with high-resolution whole-slide images; nevertheless, other fields, such as nuclear medicine and ophthalmology, are still in their infancy. Custom SSL techniques built around their unique visual characteristics could be advantageous for these.

9 Where We're Going?

There are significant gaps despite significant progress. One issue is that research uses different datasets, designs, and metrics, which makes comparisons difficult. We need standardized benchmarks and reproducible pipelines to fairly evaluate SSL approaches across different jobs and formats. Another unresolved question is whether models developed on one type of data, like CT, can be applied to other types, such as MRI. Changing between modalities could greatly lower labeling costs. Finally, explainability is still a challenge. Even though SSL improves performance, it is often unclear which parameters the model relies on. Combining explainable AI with SSL could enhance the models' use in therapeutic settings and build trust.

10 Conclusion and Future Work

Self-supervised learning is showing promising results in general and particularly in the medical imaging domain. Recently it has been blooming in all domains but more testing and organised comparisons are needed to quantify which approach performs best and under what circumstances. Also the design of Pretext tasks is another key factor which needs to be studied. Which Pretext suits which domain more is a real challenge. For example, a small crop can make a significant change in the medical images so careful design of tasks is something that needs to be figured out. Which strategy works best and why is an important aspect to be studied. Also, some special tuning, such as applying

the steady patterns in body structures, interpreting over-time changes of cardiac or fMRI data, or even integrating various kinds of information, such as radiology reports and patient information, can facilitate the learning process and result in more clinically relevant representations. These approaches not only tackle the current challenges which exist in our ecosystem but also provide a way to develop adaptable and scalable frameworks. If we could summarise the key takeaways that were taken while carrying out this research work it would be as follows:

- SSL is a promising framework in reducing the dependency on the annotated data, however a careful design of Pretext tasks is mandatory and offers a wide scope of research, especially in medical image domains.
- Different Pretext tasks have strengths and shortcomings, combining them can help leverage the strengths and be fruitful in achieving SOTA results.
- More reproducible pipelines, standardised benchmarks and Explainability of SSL frameworks remains a challenge, to fairly evaluate and use SSL in different domains.

Funding

None.

Conflict of Interest

The authors declare that there is no potential conflict of interest in this publication.

References

- [1] S.-C. Huang, A. Pareek, M. Jensen, M. P. Lungren, S. Yeung, and A. S. Chaudhari, "Self-supervised learning for medical image classification: a systematic review and implementation guidelines," *NPJ Digital Medicine*, vol. 6, no. 1, p. 74, 2023.
- [2] J. Gui, T. Chen, J. Zhang, Q. Cao, Z. Sun, H. Luo, and D. Tao, "A survey on self-supervised learning: Algorithms, applications, and future trends," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 12, pp. 9052–9071, 2024.
- [3] S. Azizi, B. Mustafa, F. Ryan, Z. Beaver, J. Freyberg, J. Deaton, A. Loh, A. Karthikesalingam, S. Kornblith, T. Chen *et al.*, "Big self-supervised models advance medical image classification," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 3478–3488.
- [4] V. R. de Sa, "Learning classification with unlabeled data," in *Advances in Neural Information Processing Systems*. MIT Press, 1994, pp. 112–119.
- [5] L. Jing and Y. Tian, "Self-supervised visual feature learning with deep neural networks: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 11, pp. 4037–4058, 2021.
- [6] J. Gui, Z. Sun, Y. Wen, D. Tao, and J. Ye, "A review on generative adversarial networks: Algorithms, theory, and applications," *IEEE Transactions on Knowledge and Data Engineering*, 2022.
- [7] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in *International conference on machine learning*. Pmlr, 2020, pp. 1597–1607.
- [8] Y. Xie, L. Gu, T. Harada, J. Zhang, Y. Xia, and Q. Wu, "Rethinking masked image modelling for medical image representation," *Medical Image Analysis*, vol. 98, p. 103304, 2024.
- [9] K. Fırlıdak, G. Çelik, and M. F. Talu, "Simclr-based self-supervised learning approach for limited brain mri and unlabeled images," *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, vol. 13, no. 4, pp. 1304–1313, 2024.
- [10] B. Dufumier, P. Gori, J. Victor, A. Grigis, M. Wessa, P. Brambilla, P. Favre, M. Polosan, C. McDonald, C. M. Piguet *et al.*, "Contrastive learning with continuous proxy meta-data for 3d mri classification," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2021, pp. 58–68.
- [11] B. Li, Y. Li, and K. W. Eliceiri, "Dual-stream multiple instance learning network for whole slide image classification with self-supervised contrastive learning," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 14 318–14 328.

- [12] S. Manna, S. Bhattacharya, and U. Pal, "Interpretive self-supervised pre-training: boosting performance on visual medical data," in *Proceedings of the Twelfth Indian Conference on Computer Vision, Graphics and Image Processing*, 2021, pp. 1–9.
- [13] A. Fedorov *et al.*, "Tasting the cake: Evaluating self-supervised generalization on out-of-distribution multimodal mri data," in *RobustML Workshop, ICLR 2021*, 2021.
- [14] J. Ouyang *et al.*, "Self-supervised longitudinal neighbourhood embedding," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*. Springer, 2021, pp. 80–89.
- [15] Q. Zhao, Z. Liu, E. Adeli, and K. M. Pohl, "Longitudinal self-supervised learning," *Medical image analysis*, vol. 71, p. 102051, 2021.
- [16] J. Osin *et al.*, "Learning personal representations from fmri by predicting neurofeedback performance," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*. Springer, 2020, pp. 469–478.
- [17] A. Fedorov *et al.*, "On self-supervised multi-modal representation learning: An application to alzheimer's disease," in *IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, 2021.
- [18] Y. Hashimoto, Y. Ogata, M. Honda, and Y. Yamashita, "Deep feature extraction for resting-state functional mri by self-supervised learning and application to schizophrenia diagnosis," *Frontiers in Neuroscience*, vol. 15, p. 696853, 2021.
- [19] F. T. Dezaki *et al.*, "Echo-rhythm net: Semi-supervised learning for automatic detection of atrial fibrillation in echocardiography," in *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2021, pp. 110–113.
- [20] Y. Zhu, "Self-supervised learning for small shot covid-19 classification," in *2021 3rd International Conference on Information Technology and Computer Communications*. IEEE, 2021, pp. 36–40.
- [21] H. Liu, D. Wei, D. Lu, J. Sun, L. Wang, and Y. Zheng, "M3ae: Multimodal representation learning for brain tumor segmentation with missing modalities," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, no. 2, June 2023, pp. 1657–1665.
- [22] F. Haghighi, M. R. H. Taher, Z. Zhou, M. B. Gotway, and J. Liang, "Transferable visual words: Exploiting the semantics of anatomical patterns for self-supervised learning," *IEEE Transactions on Medical Imaging*, 2021.
- [23] W.-T. Hsieh, J. Lefort-Besnard, H.-C. Yang, L.-W. Kuo, and C.-C. Lee, "Behavior score-embedded brain encoder network for improved classification of alzheimer disease using resting state fmri," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2020)*. IEEE, 2020, pp. 5486–5489.
- [24] S. Manna, S. Bhattacharya, and U. Pal, "Self-supervised representation learning for detection of acl tear injury in knee mr videos," *Pattern Recognition Letters*, vol. 154, pp. 37–43, 2022.
- [25] X. Wang, Y. Fang, Q. Wang, P. T. Yap, H. Zhu, and M. Liu, "Self-supervised graph contrastive learning with diffusion augmentation for functional mri analysis and brain disorder detection," *Medical Image Analysis*, vol. 101, p. 103403, 2025.