

## REVIEW

Artificial Intelligence and Applications  
2025, Vol. 00(00) 1-34  
DOI: [10.47852/bonviewAIA52026392](https://doi.org/10.47852/bonviewAIA52026392)

# Utility of Deep Learning to Address Missing Modalities from Multi-Modal Medical Imaging Studies: A Systematic Review

Jinzhao Qian<sup>1,2,†</sup> , Ankita Joshi<sup>1,†</sup> , Hailong Li<sup>1,3,4</sup>, Nehal A. Parikh<sup>5,6</sup>, Jonanthan R. Dillman<sup>1,3,4</sup>, and Lili He<sup>1,2,3,4,7,8,\*</sup><sup>1</sup> Imaging Research Center, Cincinnati Children's Hospital Medical Center, USA<sup>2</sup> Department of Computer Science, University of Cincinnati, USA<sup>3</sup> Department of Radiology, Cincinnati Children's Hospital Medical Center, USA<sup>4</sup> Department of Radiology, University of Cincinnati College of Medicine, USA<sup>5</sup> Perinatal Institute, Cincinnati Children's Hospital Medical Center, USA<sup>6</sup> Department of Pediatrics, University of Cincinnati College of Medicine, USA<sup>7</sup> Department of Biomedical Engineering, University of Cincinnati College of Medicine, USA<sup>8</sup> Department of Biomedical Informatics, University of Cincinnati College of Medicine, USA

**Abstract:** Missing modalities pose a significant challenge on multi-modal studies by disrupting the comprehensive analysis of diverse data sources. Deep learning addresses this issue by employing algorithms that can effectively infer and integrate the absent information, thereby ensuring robustness and accuracy of the models while increasing the study's statistical power. This study aims to provide a systematic literature review on deep learning solutions for missing imaging modalities in multi-modal medical data analysis. Articles on PubMed, IEEE explore digital library, and Scopus were searched in the range from January 2013 to May 2025. This systematic search and review identified 234 articles. Adhering to the specified search criteria, 61 published studies were eligible. Among these, 47% employed image synthesis methods, 20% applied knowledge transfer methods, and 33% used latent feature space-based methods. The paper explores the research gaps and challenges associated within each of these categories. Additionally, this review paper illuminates the popular public datasets for multi-modal studies with missing modalities. Furthermore, it presents evaluation metrics and their key attributes. The review concludes with its limitations and a detailed discussion of current challenges and future directions in this domain.

**Keywords:** deep learning, missing modalities, image synthesis, knowledge transfer, latent space, medical image analysis

## 1. Introduction

Over recent decades, medical advances have made it possible to generate abundant medical data for individual patients from various resources (e.g., imaging systems, blood tests, and histopathology data). Medical data obtained from different resources are referred to as multi-modal data. Each data modality represents a different type of data, such as images from medical imaging scans, text from electronic health records (EHR), histopathologic data from tissue specimens, or genetic data [1]. Advanced neural network methods have been applied to interpret these modalities of medical data generated from diverse sources. These methods aim to leverage the complementary information presented in different modalities while also utilizing the unique strengths of each modality to improve model performance and enhance understanding of complex systems [2]. Multi-modal methods have been applied across various clinical domains [3]. Combining multimodal information

enables healthcare professionals to develop a holistic view of a patient's health condition, aiding in diagnosis, treatment planning, and/or therapy monitoring [4]. For instance, in neuroimaging, the integration of structural Magnetic Resonance Imaging (MRI) with functional MRI and diffusion MRI provides detailed information about brain structure, function, and microstructure, enhancing the management of neurological disorders [5]. Similarly, in oncology, combining data from Computed Tomography (CT), MRI, and Positron Emission Tomography (PET) scans improves tumor analysis and treatment response and assessment, enabling personalized cancer care [6]. Furthermore, multi-modal methods drive medical research by exploring complex biological processes, leading to innovative diagnostic and therapeutic advances [2]. Multi-modal data fusion algorithms can fully take advantage of informative and abundant features in cancer survival prediction [7], while multi-modal ensemble techniques are suitable for decision making in medical datasets [8].

One of the most significant challenges to the adoption of multi-modality is the high prevalence of missing data. In real-world clinical settings, it is not uncommon for certain imaging modalities to be unavailable due to equipment limitations, patient-related factors, or technical issues. Additionally, artifacts or inadequate capture of specific

\*Corresponding author: Lili He, Department of Radiology, Cincinnati Children's Hospital Medical Center, University of Cincinnati College of Medicine, USA. Email: [Lili.He@cchmc.org](mailto:Lili.He@cchmc.org)

†Co-first author

regions of interest within an image further complicates the analysis process. Traditional methods for dealing with missing data often entail strategies such as complete case analysis, which assumes that the data are free of missing data and only uses cases with complete datasets for analysis [9], or simplistic imputation techniques such as mean substitution or last observation carried forward [10]. Nevertheless, these methodologies exhibit significant shortcomings. Complete case analysis reduces the study's statistical power and may lead to biased results if data are not missing at random, while simple imputation methods may introduce inaccuracies and underestimate variability, affecting the reliability of subsequent analyses. The assumption that the dataset has no missing data leads to models only working well theoretically and cannot be adapted to real-world scenarios. Moreover, these approaches fail to fully utilize available information, potentially limiting the insights gained from multi-modal data integration. Whereas, in deep learning-based approaches missing modalities will result in an insufficient dataset for training [11]. This leads to serious problems of decreased model accuracy, potential biases, and unreliable predictions, ultimately compromising the effectiveness of the model in clinical applications.

Deep learning methods have provided an alternative solution to the missing modality problem [12]. Their inherent capability to discern intricate patterns in data renders it highly suitable for managing the complexities of high-dimensional and heterogeneous multiple modalities, eliminating the need for manual feature engineering [13]. They can automatically extract and integrate relevant features from multimodal MRI, CT, and PET scans for accurate disease diagnosis without the need for manual feature extraction [14]. Besides, their adaptability ensures precise imputation of missing data, with the utilization of extensive datasets bolstering generalization across varied clinical contexts. For example, in the EHR, deep learning models exhibit proficiency in facilitating more accurate imputation by effectively learning intricate patterns and correlations within the patients' data, which enables precise predictions and the customization of treatments to the unique needs of individual patients [15]. Furthermore, they adeptly incorporate contextual cues and spatial relationships within multiple modalities, enhancing data fidelity and facilitating thorough analyses. For instance, in pathology slides, deep learning models can simultaneously analyze histopathological images alongside clinical data and genetic information to deliver a thorough understanding of disease progression and treatment response [16].

The past few years have witnessed heightened focus on utilizing deep learning methods to address the issue of missing data in multi-modal MRI-based analysis, with published surveys discussing the missing modality issue of MRI data [17, 18]. However, these surveys typically group methods based on technical implementation details or specific modalities and mainly focus on the brain tumor segmentation task, which makes cross-task generalization less clear. By contrast, our framework emphasizes the underlying strategy for handling missing modalities, providing a broader and more unifying perspective. Therefore, in this work, we provide a systematic literature review focusing on all types of imaging modalities and applicable to all tasks within the medical domain like image quality analysis, classification, prediction, and segmentation. Such a review will furnish invaluable insights into the available solutions, limitations, and optimal practices associated with deploying deep learning strategies for addressing missing data issues in multi-modal medical data analysis. In this review, we summarize the deep learning-based methods which address the missing modality problem. In that, we target methods that specifically provide solutions for missing images. We aim to cover every contributory method in the last decade. This survey contributes to the following:

- 1) Systematically review missing modality deep learning techniques in the last 10 years which have provided solutions to missing modality issues in the context of multi-modal medical image analysis.

- 2) Categorize the identified solutions based on their approach and present their benefits and limitations.
- 3) Summarize datasets and assessment criteria applied in the studies.
- 4) Analyze the main findings and outline future directions for research in this domain.

The remainder of the review is organized as follows: Section 2 outlines the review planning; Section 3 presents key findings; Section 4 discusses and interprets the results; Section 5 states limitations in our research; Section 6 highlights future directions; and Section 7 makes the conclusion of the review.

## 2. Review Planning

This section focuses on organizing the review: clearly specifying the research questions relevant to the study, detailing the information sources, and outlining the inclusion criteria.

### 2.1. Key research questions

- 1) What approach is proposed to address the missing modality problem?
- 2) Which dataset is used?
- 3) Under which methodological category does the proposed approach fall?
- 4) What pre-processing approaches are used?
- 5) What criteria are used for evaluation?
- 6) Is the code publicly available?
- 7) Does the article carry out external validation?

### 2.2. Data sources and search strategy

Our study followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [19]. We searched both general databases (Scopus, PubMed) and a subject-specific database (IEEE Xplore) to ensure broad biomedical coverage while also capturing technical contributions often published in engineering venues. The search strategy was developed for the purpose of reviewing deep learning methods for missing modalities in multimodal medical imaging. Queries combined terms related to multimodality, missing data, clinical tasks (e.g., classification and segmentation), medical imaging context, and deep learning techniques. This blended approach allowed us to balance breadth and depth and ensured that both methodological and application-oriented studies were included. The queries used for the respective sources of publications spanning from January 2013 to May 2025 have been listed in Table 1.

### 2.3. Inclusion criteria

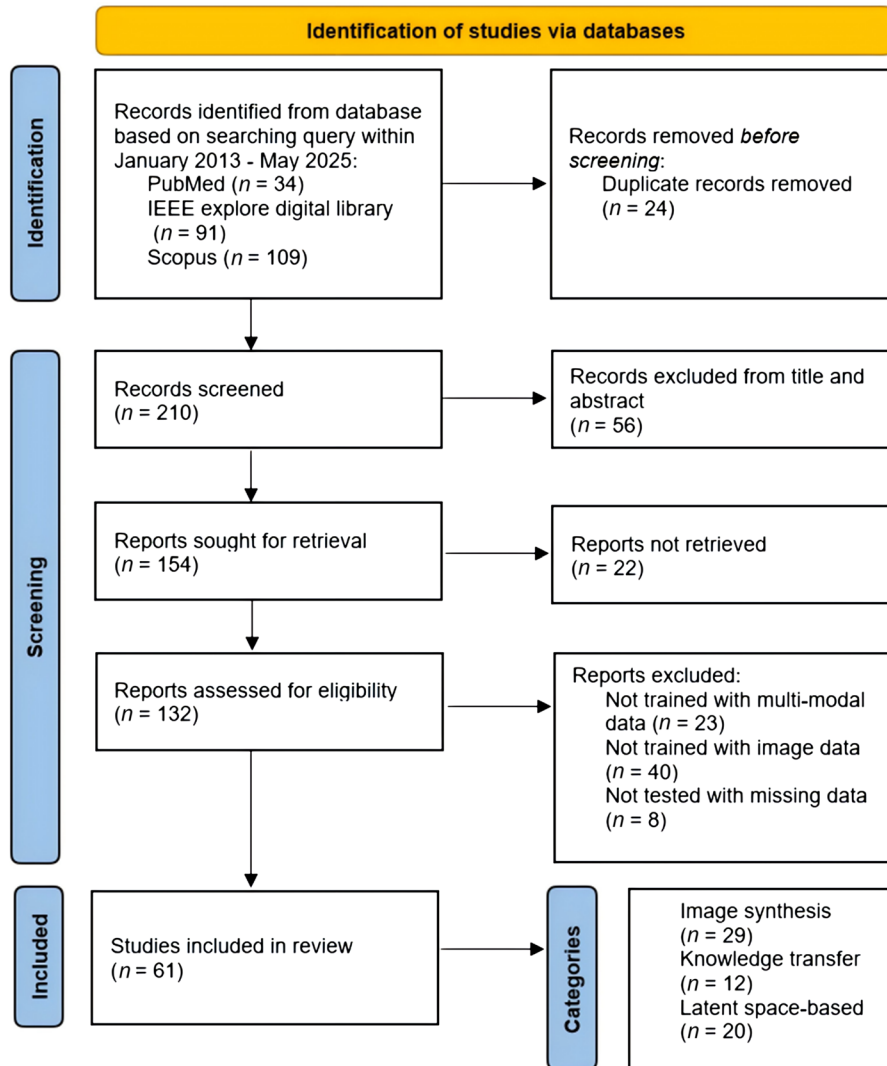
Inclusion requirements were: (a) original research article; (b) published within January 2013 and May 2025; (c) published in English language; (d) required to be a multi-modal study involving different modalities of medical images, or imaging and any other non-imaging modality; (e) proposing solutions to missing imaging modality; (f) proposing deep learning-based solutions; and (g) applied to any classification, prediction, or segmentation task in medical domain.

Specifically, papers' titles and abstracts were sought on each of the previously mentioned information sources using the respective search queries as given in Table 1. Figure 1 presents the flow diagram based on the PRISMA guidelines. In total, 234 records were initially identified. Following the first exclusion phase, 210 remained for abstract screening. Of these, 132 met the criteria for full-text review. Finally, 61 studies were retained for this review, with 29 image synthesis methods, 12 knowledge transfer methods, and 20 latent feature space-based methods. The resulting 61 studies were cross-checked by a second

Table 1  
Search query

| Source of information                     | Query used                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|-------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| PubMed <sup>1</sup>                       | ((multimodal) OR (fusion) OR (ensemble) OR (feature integration) OR (multiparametric)) AND ((missing data[Title/Abstract]) OR (missing modality[Title/Abstract]) OR (incomplete data[Title/Abstract])) AND ((classification [Title/Abstract]) OR (regression[Title/Abstract]) OR (prediction[Title/Abstract]) OR (diagnosis[Title/Abstract]) OR (segmentation[Title/Abstract])) AND ((medical image) OR (medical application)) AND (deep learning)                                                                                                                                                                                            |
| IEEE explore digital library <sup>2</sup> |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| Scopus <sup>3</sup>                       | ( ALL ( multimodal ) OR ALL ( fusion ) OR ALL ( ensemble ) OR ALL ( feature_integration ) OR ALL ( multi-parametric ) )<br>AND ( TITLE-ABS-KEY ( missing AND data ) OR TITLE-ABS-KEY ( missing AND modality ) OR TITLE-ABS-KEY ( incomplete AND data ) )<br>AND ( TITLE-ABS-KEY ( classification ) OR TITLE-ABS-KEY ( regression ) OR TITLE-ABS-KEY ( prediction ) OR TITLE-ABS-KEY ( diagnosis ) OR TITLE-ABS-KEY ( segmentation ) )<br>AND ( ALL ( medical_image ) OR ALL ( medical_application ) )<br>AND ( ALL ( deep_learning ) )<br>AND ( LIMIT-TO ( SUBJAREA,"COMP" ) ) AND ( LIMIT-TO ( DOCTYPE,"ar" ) OR LIMIT-TO ( DOCTYPE,"cp" ) ) |

Figure 1  
Flow diagram for retrieved articles based on the PRISMA guidelines



<sup>1</sup>(<https://pubmed.ncbi.nlm.nih.gov/>)

<sup>2</sup>(<https://ieeexplore.ieee.org/Xplore/home.jsp>)

<sup>3</sup>(<https://www.scopus.com/>)

reviewer, and disagreements were resolved through team discussion until consensus was reached.

## 2.4. Data extraction

For each included study, we extracted key information using a structured template, covering publication year, dataset(s), imaging modalities, methodological approach, preprocessing steps, model architecture, evaluation metrics, performance results, code availability, and use of external validation. Data extraction was carried out by one reviewer and verified by a second. To synthesize findings, studies were grouped into three methodological categories (image synthesis, knowledge transfer, and latent feature space-based methods). Within each category, results were compared in terms of performance, robustness to missing modalities, reproducibility (code availability, external validation), and task type. Evidence was summarized narratively and with structured tables (Tables 4–6), alongside a standardized comparison table (Table 8) highlighting core attributes such as performance, robustness, and reproducibility.

## 3. Results

With above search strategy, we retrieved 234 articles published between January 2013 and May 2025. After manually removing articles that did not meet the inclusion criteria and searching the bibliography of eligible articles, 61 articles met the inclusion criteria. The results are presented as follows: First, we outline the datasets utilized across the articles included in this review. Then, we present a compilation of the 61 research articles, organized based on the established categories and finally we summarize the evaluation metrics used by these studies.

### 3.1. Dataset

We list all the public multimodal datasets used by the 61 articles included in this review in Table 2. For each dataset we include the online location to access the dataset, the dataset details in terms of number of subjects, and the modalities available with the dataset.

### 3.2. Deep learning methods with missing imaging modality

The purpose of this section is to investigate the questions raised in the review planning process in Section 2. Without considering specific tasks (e.g., classification, regression, or segmentation), multi-modal deep learning models take multiple modalities of imaging data, and extract the high-level feature representations (i.e., feature embeddings).

In this work, we categorized all reviewed deep learning methods into (i) image synthesis methods, (ii) knowledge transfer methods, and (iii) latent feature space-based methods. Figure 2 illustrates the categories presented in this review. Briefly, image synthesis methods synthesize the missing imaging modalities with existing ones, knowledge transfer methods focus on establishing new architectures/deep learning models to transfer knowledge, and latent feature space-based methods work in the shared feature sub-space. We present all 61 articles identified in this review based on these three categories. Table 3 presents all identified papers under each category, with 29 studies in image synthesis methods, 12 studies in knowledge transfer methods and 20 studies in latent feature space-based methods. Figure 3 summarizes their distribution in a taxonomy diagram. In Tables 4–6, for each category of methods, we provided the preprocessing pipeline of the method, the basic architecture of the deep learning model, the main context of the proposed solution for handling missing modality, the medical task performed, and the evaluation metrics used to assess the performances. To show the accessibility of the methods, we also

mention code availability of the proposed solution and the use (or absence) of external validation.

#### 3.2.1. Image synthesis methods

Image synthesis methods aim at recovering the missing imaging modalities from the available imaging modalities. These methods typically were developed using Generative Adversarial Networks (GAN) [97] and its variants [98]. The synthesis models consist of two subnetworks: one generator that synthesizes the missing modalities from available ones, and one discriminator that judges whether the input image is real or synthesized by generator [99]. These subnetworks are trained simultaneously under an adversarial loss. As a result, the generator's performance is improved to create satisfying images as compensation to missing ones. To improve the quality of synthesized image, multi-modal learning techniques are applied in the generator to correlate features from multiple modalities [40]. The synthesized images can either serve as supplementary training data to help the downstream models better understand the tasks or be used to augment training data, in order to improve the models' robustness and generalization capabilities.

Despite the obvious advantage of recovering missing imaging modality information directly, the training cost of GAN is relatively high due to their model complexity. What is more, the training of GAN will become unstable when encountering issues like mode collapse [100], resulting in poor synthesis performance. Table 4 gives a detailed list of all studies included in this review that use image synthesis approaches.

#### 3.2.2. Knowledge transfer methods

Knowledge transfer methods first develop a source model that encodes all data modalities. The representations learned from the source model are then transferred to the target model, which has incomplete modalities as input. This target model is then utilized for the downstream task [68]. There are two ways to transfer the learned modalities from the source model to the target model, namely, using the knowledge distillation method [65], or by using domain adaptation [66].

Knowledge distillation involves training the target model in accordance with the source model. First, the source model with a larger number of parameters is trained with the full modalities until it achieves high performance on the target task. The target model of less complexity learns from the source model by capturing the encoded knowledge in its outputs and then is trained with incomplete datasets. The forward propagation is performed on both models while the back propagation is only performed on the target model. The output of the source model can be used as soft labels to guide the training of target model [103]. The discrepancy between the target model's predictions and the soft labels of the source model is measured by the target's distillation loss, which is combined with traditional learning loss in the training of the target model.

Domain adaptation methods are typically used in transfer learning approaches for handling missing modalities to uncover the common latent features across the source and target domains. The source model encodes complete modalities, whereas the target model encodes the incomplete modalities separately, and features are extracted from both models. A discriminator is used to bridge the domain gap between the output of the source and target model. A similarity loss is used with the discriminator to map the features to a similar distribution, and a consistency loss is applied to minimize the distance between the distributions [104].

Knowledge transfer methods outperform other non-synthesizing methods in multiple conditions of missing modalities and are capable of transferring to other tasks [67]. Nevertheless, it could be affected by limited availability or imbalanced datasets, resulting in difficulty

**Table 2**  
**Summary of the datasets used in the papers included in the review**

| Multi-modal datasets                                               |                                                                                                                                             |                                                                                                                                                                                                                                                                                                                  |                                                                                                                         |
|--------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------|
| Dataset                                                            | Online link                                                                                                                                 | Subjects                                                                                                                                                                                                                                                                                                         | Modalities                                                                                                              |
| Alzheimer's Disease NeuroImaging Initiative (ADNI) [20]            | <a href="https://adni.loni.usc.edu/">https://adni.loni.usc.edu/</a>                                                                         | 200 elderly cognitively normal, 400 Mild Cognitive Impairment (MCI), and 200 Alzheimer's disease (AD) subjects                                                                                                                                                                                                   | Clinical, genetic, MRI, PET, biospecimen                                                                                |
| Brain Tumor Segmentation Challenge (BraTS (2021)) [21]             | <a href="http://braintumorsegmentation.org/">http://braintumorsegmentation.org/</a>                                                         | 2040 glioma cases (8000 Multiparametric MRI scans)                                                                                                                                                                                                                                                               | MRI modalities: T1, T1 post-contrast (T1c), T2, T2 Fluid-Attenuated Inversion Recovery (FLAIR)                          |
| Ischemic Stroke Lesion Segmentation (ISLES) [22]                   | <a href="https://www.smir.ch/ISLES/Start2015">https://www.smir.ch/ISLES/Start2015</a>                                                       | 36 cases on subtask: sub-acute ischemic stroke lesion segmentation (SISS)                                                                                                                                                                                                                                        | MRI modalities: T1, T1c, T2, T2 FLAIR, Diffusion-weighted Imaging (DWI)                                                 |
| IXI Brain Development                                              | <a href="https://brain-development.org/ixi-dataset/">https://brain-development.org/ixi-dataset/</a>                                         | Nearly 600 scans from normal, healthy subjects                                                                                                                                                                                                                                                                   | MRI modalities: T1, T2, Proton Density (PD), Magnetic Resonance Angiography (MRA), DWI                                  |
| Open Access Series of Imaging Studies (OASIS-3) [23]               | <a href="https://sites.wustl.edu/oasisbrains/home/oasis-3/">https://sites.wustl.edu/oasisbrains/home/oasis-3/</a>                           | 1378 participants collected across several ongoing projects through the WUSTL Knight ADRC over the course of 30years, include 755 cognitively normal adults and 622 individuals at various stages of cognitive decline ranging in age from 42-95yrs.                                                             | MRI modalities: T1, T2, T2FLAIR, etc.<br>PET, CT                                                                        |
| The Cancer Genome Atlas Program (TCGA)                             | <a href="https://www.cancer.gov/ccg/research/genome-sequencing/tcga">https://www.cancer.gov/ccg/research/genome-sequencing/tcga</a>         | Including Kidney Clear Cell Carcinoma (KIRC) (385 cases), Liver Hepatocellular Carcinoma (LIHC) (287 cases), Esophageal Carcinoma (ESCA) (153 cases), Lung Squamous Cell Carcinoma (LUSC) (438 cases), Lung Adenocarcinoma (LUAD) (452 cases), and Uterine Corpus Endometrial Carcinoma (UCEC) (387 cases), etc. | Genomic data, MRI, Region Of Interest (ROI) images, diagnostic Whole Slide Imaging (WSI) images, clinical records, etc. |
| MIMIC Chest X-ray (MIMIC-CXR) [24]                                 | <a href="https://physionet.org/content/mimic-cxr/2.0.0/">https://physionet.org/content/mimic-cxr/2.0.0/</a>                                 | 65,379 patients presenting to the Beth Israel Deaconess Medical Center Emergency Department between 2011–2016                                                                                                                                                                                                    | Clinical time series data, X-ray                                                                                        |
| Parkinson's Progression Markers Initiative (PPMI) [25]             | <a href="https://www.ppmi-info.org/access-data-specimens/download-data/">https://www.ppmi-info.org/access-data-specimens/download-data/</a> | 400 recently diagnosed Parkinson Disease and 200 healthy subjects                                                                                                                                                                                                                                                | T1 MRI, Diffusion Tensor Imaging (DTI) images, Single Nucleotide Polymorphism (SNP)                                     |
| LONI Probabilistic Brain Atlas (LPBA40) [26]                       | <a href="https://loni.usc.edu/research/atlas">https://loni.usc.edu/research/atlas</a>                                                       | 40 healthy, normal subjects                                                                                                                                                                                                                                                                                      | T1 MRI                                                                                                                  |
| Synapse [27]                                                       | <a href="https://www.synapse.org/#!Synapse:syn3193805/wiki/217753">https://www.synapse.org/#!Synapse:syn3193805/wiki/217753</a>             | 34 subjects for MRI and 30 subjects for CT                                                                                                                                                                                                                                                                       | MRI, CT                                                                                                                 |
| Combined (CT-MR) Healthy Abdominal Organ Segmentation (CHAOS) [28] | <a href="https://chaos.grand-challenge.org/">https://chaos.grand-challenge.org/</a>                                                         | 80 patients, 40 of them (22 males, 18 females, ages between 18 and 63) went through a single CT scan and 40 of them (23 male, 17 female, ages between 18 and 76) went through MR scans                                                                                                                           | MRI, CT                                                                                                                 |
| Chinese Brain Molecular and Functional Mapping (CBMFM) [29]        | Upon request                                                                                                                                | 646 subjects from 18 to 82 years old collected from four medical centers                                                                                                                                                                                                                                         | MRI modalities: T1, T2, T2FLAIR                                                                                         |

Table 2  
(Continued)

| Multi-modal datasets                                          |                                                                         |                                                                                                                                                   |                                                                |
|---------------------------------------------------------------|-------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------|
| Dataset                                                       | Online link                                                             | Subjects                                                                                                                                          | Modalities                                                     |
| Early life adversity biological embedding (eLABE)             | Upon request                                                            | 127 neonates (postmenstrual age=41.1 ± 1.5 weeks, female N = 59, white N = 42)                                                                    | MRI modalities: T1, T2, resting state functional MRI (rs-fMRI) |
| Environmental influences on child health outcomes (ECHO) [30] | Upon request                                                            | including 10 infants (age=41.2 ± 1.9 weeks, female N = 5, white N = 8)                                                                            | MRI modalities: T1, T2, rs-fMRI                                |
| Multiple Sclerosis with the MS Grand Challenge (MSGC) [31]    | <a href="http://www.ia.unc.edu/MSseg/">http://www.ia.unc.edu/MSseg/</a> | 45 cases, 25 from the Boston Children's Hospital (CHB) and 20 from the University of North Carolina (UNC)                                         | MRI modalities: T1, T2, T2FLAIR                                |
| Relapsing Remitting Multiple Sclerosis (RRMS)                 | Upon request                                                            | 300 RRMS patients (mean age = 37.5, surface distance (SD) = 10.0)                                                                                 | MRI modalities: T1, T2, T2FLAIR, T1c                           |
| Chinese Longitudinal Aging Study (CLAS) [32]                  | Upon request                                                            | 1068 elderly Chinese (42.2% male), mean age of 72.8 years (SD=8.5) completed a comprehensive cognitive, psychosocial and mental health assessment | T1 MRI                                                         |
| The Australian Imaging, Biomarkers and Lifestyle (AIBL) [33]  | Upon request                                                            | 1112 subjects, 211 AD, 133 MCI, 768 healthy controls                                                                                              | T1 MRI, PET                                                    |
| UNC/UMN Baby Connectome Project [34]                          | Upon request                                                            | 500 typically developing infants, toddlers, and preschool-aged children between birth and 5 years of age                                          | MRI modalities: T1, T2, DWI, rs-fMRI                           |
| MIDAS [35]                                                    | Upon request                                                            | 34 healthy subjects, ranging in age from 19 to 72 and of both sexes, 27 tumor cases included thirty lesions                                       | MRI modalities: T1, T2, MRA                                    |

in training or a decrease in performance [71]. Table 5 gives a comprehensive list of all the studies covered in this review that use knowledge transfer approaches.

### 3.2.3. Latent feature space-based methods

Latent feature space-based methods offer a typical solution to handling missing modalities by focusing on deriving meaningful features from existing modalities alone, bypassing intermediate processes such as imputation [106] or synthesis [107]. Latent feature space-based methods involve independently learning and embedding of the input image into a latent space for each modality. Each of the modalities is encoded respectively to extract its specific feature representations. Subsequently, these modalities are translated into a shared latent space. Fusion strategies such as mean and variance operations are utilized to establish one common latent embedding for all modalities, so that missing features can be recovered [96]. Once the shared representation is obtained, a decoder is used to reconstruct the data from it to perform the downstream tasks. Such methods are robust to missing imaging modality by learning a shared representation of the available modalities [81].

However, projecting data into a latent space can lead to information loss from the original modalities. This loss can affect the overall performance of the model, particularly if the missing modalities contain critical information [78]. Also, the learned latent vectors of

individual modalities from independent encoders are fused to create a shared representation. However, this assumes that each modality is equally informative and contributes equally to the final task, which may not always be true. Besides, interpreting the results of models based on latent feature space methods can be challenging, especially when the relationships between modalities are complex or non-linear. Table 6 gives a detailed list of all the studies included in this review that use latent feature space-based approaches.

## 3.3. Evaluation metrics

In this section, we review the evaluation metrics applied for validating the performances of the methods in missing modalities scenarios. In classification tasks, accuracy, sensitivity, specificity, and area under the curve (AUC) are the most popular choices for assessment. Dice similarity coefficient (DSC) is widely used for measuring the segmentation results, while peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM) intuitively display the quality of synthesized images of missing modalities. Table 7 gives a list of all the evaluation metrics used in the articles included in this review with the following attributes for each: the name and abbreviation of the metric, the formulation to compute the metric, and the usage of the metric.

Figure 2

Illustration of methods by category. (a) With complete modalities, multimodal images are directly used for feature extraction. When modalities are missing, three strategies are applied: (b) synthesis methods generate absent modalities from available ones, (c) knowledge transfer leverages models trained on complete data to guide models trained on partial data, (d) latent space methods fuse embeddings into shared representations

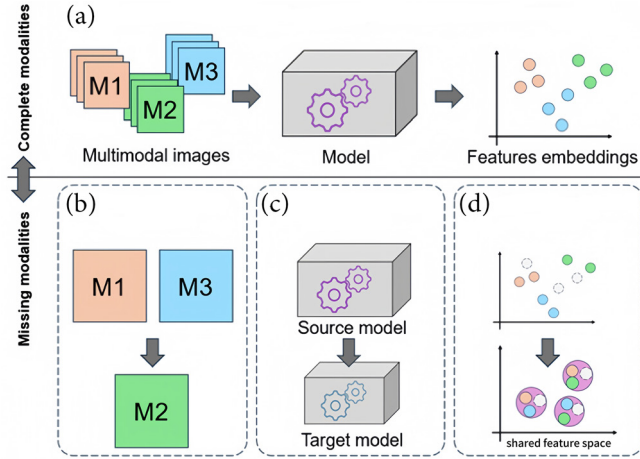


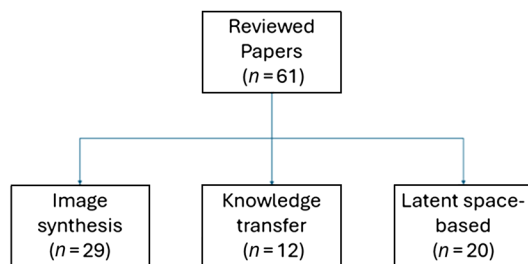
Table 3

Methodological categories identified in this review and their corresponding studies

| Category                   | References                                                                                                                                                                   | Total number |
|----------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|
| Image synthesis methods    | [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53], [54], [55], [56], [57], [58], [59], [60], [61], [62], [63], [64] | 29           |
| Knowledge transfer methods | [65], [66], [67], [68], [69], [70], [71], [72], [73], [74], [75], [76]                                                                                                       | 12           |
| Latent space-based methods | [77], [78], [79], [80], [81], [82], [83], [84], [85], [86], [87], [88], [89], [90], [91], [92], [93], [94], [95], [96]                                                       | 20           |

Figure 3

Taxonomy diagram for retrieved articles in this study



## 4. Discussion

This paper offers a systematic review of deep learning methods dealing with missing imaging modalities in medical analysis in terms

of the most popular public datasets used, the different types of deep learning approaches developed, and the most common metrics for evaluating the performance of the approaches. Though prior works have reviewed deep learning approaches for brain tumor segmentation tasks under the condition of missing MRI modalities, a systematic review that includes deep learning methods under the circumstance of missing medical imaging modalities regardless of the tasks to be performed, has not been published yet. We identified 61 papers in this review that focus on deep learning-based solutions that we categorized into image synthesis methods, knowledge transfer methods, and latent feature space-based methods. Unlike bibliometric systematic reviews, which primarily analyze publication volume, author networks, or keyword co-occurrence, our review uncovers trends at the methodological level. These include shifts in the types of deep learning models employed (e.g., increasing use of U-Nets and transformers), the prevalence of specific evaluation metrics across tasks (e.g., SSIM for image synthesis, DSC for segmentation, and AUC for classification), and reproducibility practices (e.g., proportion of studies with external validation or public code). We therefore provide a systematic methodological synthesis rather than a bibliometric mapping, which we believe is more directly useful to the medical imaging community.

In the last decade, neural network approaches have gradually been recognized for their capabilities of tackling missing modalities in medical analysis. Overall, we find that both latent feature space-based methods favored for their straightforward deployment, and image synthesis methods due to the popularity of GAN, are the more favorable techniques for dealing with missing modalities. Knowledge transfer methods show noteworthy performance in cases where multiple modalities of the data are missing.

Firstly, our results show that 47% of the reviewed papers use image synthesis methods. 17% of articles focus on methods that are implemented for the segmentation purpose [44, 46, 55, 56, 64], 38% are implemented for the classification tasks [37, 39–41, 47, 50, 53, 57–59, 63], and 4% for the prediction use [48]. 41% of the reviewed papers under image synthesis-based methods focus on the evaluation of the quality of synthesized images [36, 38, 42, 43, 45, 49, 51, 52, 54, 60–62]. All the reviewed image synthesis methods utilize GAN or its extension framework, consisting of one or two generators and one discriminator. However, the architecture of the generator to synthesize images for the recovery of the missing modalities varies. Nearly half of the GAN-based synthesis networks apply CNN-based structure as the generator. Nevertheless, CNN-based generator has several drawbacks, such as fixed input–output channels, lack of interpretability [45], and loss of low-level spatial information [49]. U-net, which is capable of channel-wise feature fusion, wins more favor in the choice of generator. It allows region-specific feedback, enhancing the generator’s learning process and leading to superior image synthesis quality. According to our review, U-net is utilized as the generator in 36% of the papers and has been especially popular in the last three years. It is worth mentioning that 12% of the papers combine CNN or U-net with the transformer in the image synthesis network. Leveraging the attention mechanism, the transformer performs efficient and accurate synthesis by capturing cross-modal correlations within input modalities while removing redundant information, and the interpretability of the model could be increased by visualizing the attention score. In [44], an efficient generator is designed by the combination of transformer and CNN, enabling the model to have global sensitivity as well as detailed local modeling. In terms of the discriminator, almost all the papers choose CNN. Only a few papers attempt to make effective improvements on the discriminator, such as the task-induced design of the discriminator aimed at integrating image synthesis with downstream tasks [53].

Next, 20% of the reviewed papers use knowledge transfer methods, which have garnered increased attention among researchers in recent years. 58% of these methods are implemented to perform

Table 4  
Image synthesis methods dealing with missing imaging modality

| Image synthesis methods      |                                                                                                                                                                                                             |                                                                                                                                                                                                 |                             |                                                |                                            |                                                                                                                                    |      |                        |
|------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------|------------------------------------------------|--------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------|------|------------------------|
| Year/<br>author/<br>citation | Dataset/no. of subjects/<br>modalities                                                                                                                                                                      | Main context                                                                                                                                                                                    | Architecture                | Preprocessing                                  | Task                                       | Metrics/performance                                                                                                                | Code | External<br>validation |
| 2025/Kebaili,<br>et al. [36] | BraTS 2021/1251 (60%<br>training, 20% validation,<br>20% testing)/MRI (T1,<br>FLAIR, T1c and T2)                                                                                                            | Propose a diffusion-based<br>generative model that<br>studies a unified feature<br>representation over full<br>input modalities                                                                 | Variational<br>autoencoders | Resizing                                       | ① synthesized image<br>quality measurement | Mean squared error (MSE);<br>Peak signal-to-noise ratio<br>(PSNR); SSIM; Learned per-<br>ceptual image patch similarity<br>(LPIPS) | N/A  | ×                      |
|                              |                                                                                                                                                                                                             |                                                                                                                                                                                                 |                             |                                                |                                            | ① FLAIR                                                                                                                            |      |                        |
|                              |                                                                                                                                                                                                             |                                                                                                                                                                                                 |                             |                                                |                                            | 0.0012; 30.707; 0.9130;                                                                                                            |      |                        |
|                              |                                                                                                                                                                                                             |                                                                                                                                                                                                 |                             |                                                |                                            | 0.0283T1                                                                                                                           |      |                        |
|                              |                                                                                                                                                                                                             |                                                                                                                                                                                                 |                             |                                                |                                            | 0.0017; 31.514; 0.9548;                                                                                                            |      |                        |
|                              |                                                                                                                                                                                                             |                                                                                                                                                                                                 |                             |                                                |                                            | 0.0155                                                                                                                             |      |                        |
|                              |                                                                                                                                                                                                             |                                                                                                                                                                                                 |                             |                                                |                                            | T1CE                                                                                                                               |      |                        |
|                              |                                                                                                                                                                                                             |                                                                                                                                                                                                 |                             |                                                |                                            | 0.0008; 32.631; 0.9311;                                                                                                            |      |                        |
|                              |                                                                                                                                                                                                             |                                                                                                                                                                                                 |                             |                                                |                                            | 0.0265                                                                                                                             |      |                        |
|                              |                                                                                                                                                                                                             |                                                                                                                                                                                                 |                             |                                                |                                            | T2                                                                                                                                 |      |                        |
| 2025/Guo, et.al<br>[37]      | Dataset from the Second<br>Affiliated Hospital of<br>Kunming Medical Uni-<br>versity/359 (115 benign,<br>244 malignant)/ultra-<br>sound images of lesions<br>(B-mode), Color Doppler<br>Flow Imaging (CDFI) | Proposed a multimodal<br>liver tumor classifica-<br>tion framework that<br>uses generative models<br>(U-GAT-IT, MSA-GAN)<br>to reconstruct missing<br>images for cross-modal<br>supplementation | GAN                         | Annotation, histogram equal-<br>ization.       | Classification<br>① liver tumor            | Accuracy (ACC); Precision<br>(PRE); Recall ((REC); F1<br>score (F1S); AUC<br>① 88.57; 87.97; 86.32; 0.87;<br>0.95                  | N/A  | ×                      |
|                              |                                                                                                                                                                                                             |                                                                                                                                                                                                 |                             |                                                |                                            |                                                                                                                                    |      |                        |
| 2024/Deng, et al.<br>[38]    | SYSU abdominal data-<br>set/24 (1729 slices train-<br>ing, 299 slices testing/<br>MRI, CT<br>IXI/40 (30 training, 10<br>testing)/T1 MRI, T2 MRI<br>(PD)                                                     | Integrating Channel<br>Attention Module into<br>the generator for better<br>feature extraction, in-<br>corporating VGG16 and<br>multiple loss terms                                             | GAN, VGG16                  | Images registration, removal,<br>normalization | ① synthesized image<br>quality measurement | PSNR; SSIM<br>① T1toT2<br>27.084; 0.875                                                                                            | N/A  | √                      |

Table 4  
(Continued)

| Image synthesis methods               |                                                                                                                                                                                                                     |                                                                                                                                                    |                            |                                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                                                                                                   |                                                                                                                                                                                                                     |      |                        |
|---------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|------------------------|
| Year/<br>author/<br>citation          | Dataset/no. of subjects/<br>modalities                                                                                                                                                                              | Main context                                                                                                                                       | Architecture               | Preprocessing                                                                                                                                                                                                                                                                                                                                                                                         | Task                                                                                                                                                              | Metrics/performance                                                                                                                                                                                                 | Code | External<br>validation |
| 2023/Zhang,<br>Mengyi, et al.<br>[39] | ADNI/370/MRI, FDG-<br>PET                                                                                                                                                                                           | Synthesize PET images<br>using a pyramidal<br>attention mechanism and<br>standard discriminators                                                   | GAN, U-net                 | -(MRI) spatial normalization,<br>brain image segmentation,<br>gray matter image acquisition,<br>gray matter image modula-<br>tion, image smoothing<br>-(PET) head correction, regis-<br>tration and normalization,<br>PET image normalization,<br>PET image smoothing<br>-(fusion) MRI brain gray mat-<br>ter segmentation, brain gray<br>matter and metabolic informa-<br>tion fusion Classification | Classification<br>① Alzheimer's disease<br>(AD) vs. Cognitively Normal (CN)<br>② Mild Cognitive Impairment<br>(MCI) vs. CN<br>③ AD vs. MCI<br>④ AD vs. MCI vs. CN | ACC; Sensitivity (SEN); Spec-<br>ificity (SPE)<br>① missing PET<br>93.4; 97.5; 92.8<br>② missing PET<br>92.5; 92.2; 92.5<br>③ missing PET 93.1;91.5;97.3<br>④ missing PET 89.9; 82.5;<br>85.9                       | N/A  | ×                      |
| 2023/Gao,<br>Xingyu, et al.<br>[40]   | ADNI-1/821(196AD,<br>227CN,168pMCI,<br>230sMCI)/T1 MRI, T2<br>MRI, FDG-PET<br>ADNI-2/534(156AD,<br>200CN,66pMCI,112sM-<br>CI)/T1 MRI, T2 MRI,<br>FDG-PET<br>OASIS-3/345(174AD,<br>171CN)/T1 MRI, T2 MRI,<br>FDG-PET | Use a feature autore-<br>gression branch and a<br>task-level discriminator<br>with transverse blocks;<br>supervise local regions at<br>voxel-level | GAN, U-net,<br>transformer | Skull-stripping, region<br>extraction, registration, trans-<br>formation                                                                                                                                                                                                                                                                                                                              | Classification<br>① AD vs. CN<br>② progressive MCI<br>(pMCI) vs. stable MCI<br>(sMCI)                                                                             | ACC; SEN; SPE; AUC<br>① missing T2MRI and PET<br>(test on ADNI-2)<br>94.4; 93.0; 95.5; 97.6 (test on<br>OASIS-3)<br>88.4; 84.6; 92.3; 94.4<br>② missing T2MRI and PET<br>(test on ADNI-2)<br>77.8; 75.4; 79.6; 82.8 | N/A  | ✓                      |
| 2023/Wang, Ton-<br>ghui, et al. [41]  | 2Dslices of breast<br>MRI/98(2245 slices)<br>(65%training, 10%valida-<br>tion, 25%testing)/DCE-<br>MRI, DWI                                                                                                         | Utilize a GAN-like<br>network to model the<br>correlations between<br>cross-sequence feature<br>representations of differ-<br>ent modalities       | cycleGAN [98],<br>ResNet   | Manual and semi-automatic<br>segmentation, lesion enhance-<br>ment                                                                                                                                                                                                                                                                                                                                    | Classification (breast<br>tumor lesion task)<br>① slice level<br>② patient level                                                                                  | ACC; PRE; REC; FIS<br>① 0.8561; 0.8783; 0.8720;<br>0.8740<br>② 0.8966; 0.9011; 0.9540;<br>0.9256                                                                                                                    | N/A  | ×                      |

Table 4  
(Continued)

| Image synthesis methods                                            |                                                                                                                                                                                                           |                                                                                                                                                                                                                 |                        |                                                                                       |                                            |                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                                              |                        |
|--------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------|---------------------------------------------------------------------------------------|--------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------|------------------------|
| Year/<br>author/<br>citation                                       | Dataset/no. of subjects/<br>modalities                                                                                                                                                                    | Main context                                                                                                                                                                                                    | Architecture           | Preprocessing                                                                         | Task                                       | Metrics/performance                                                                                                                                                                                                                                                                                                                                                                                                          | Code                                                                                                         | External<br>validation |
| 2023/Jiang,<br>Yang, Shuang<br>Zhang, and<br>Jianning Chi.<br>[42] | BraTS 2018/285 (257<br>training, 28 testing)/MRI<br>(T1, FLAIR, T1c and T2)                                                                                                                               | Train generator in a<br>supervised manner<br>with T2 images and<br>multimodal classification<br>labels, while a multi-<br>branch convolutional<br>neural network (CNN)<br>is introduced as the<br>discriminator | GAN, U-net             | Filtering, resizing, labeling                                                         | ① synthesized image<br>quality measurement | PSNR; SSIM                                                                                                                                                                                                                                                                                                                                                                                                                   | N/A                                                                                                          | X                      |
|                                                                    |                                                                                                                                                                                                           |                                                                                                                                                                                                                 |                        |                                                                                       | ② whole tumor (WT)                         | ① T2toFLAIR<br>22.2560; 0.9478                                                                                                                                                                                                                                                                                                                                                                                               |                                                                                                              |                        |
|                                                                    |                                                                                                                                                                                                           |                                                                                                                                                                                                                 |                        |                                                                                       | ③ tumor Core (TC)                          | T2toT1<br>29.9289; 0.9713                                                                                                                                                                                                                                                                                                                                                                                                    |                                                                                                              |                        |
|                                                                    |                                                                                                                                                                                                           |                                                                                                                                                                                                                 |                        |                                                                                       | ④ enhanced tumor<br>(ET)                   | T2toT1c<br>24.6464; 0.9537                                                                                                                                                                                                                                                                                                                                                                                                   |                                                                                                              |                        |
|                                                                    |                                                                                                                                                                                                           |                                                                                                                                                                                                                 |                        |                                                                                       | DSC                                        |                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                                              |                        |
|                                                                    |                                                                                                                                                                                                           |                                                                                                                                                                                                                 |                        |                                                                                       |                                            | ② 0.8836                                                                                                                                                                                                                                                                                                                                                                                                                     |                                                                                                              |                        |
|                                                                    |                                                                                                                                                                                                           |                                                                                                                                                                                                                 |                        |                                                                                       |                                            | ③ 0.8056                                                                                                                                                                                                                                                                                                                                                                                                                     |                                                                                                              |                        |
|                                                                    |                                                                                                                                                                                                           |                                                                                                                                                                                                                 |                        |                                                                                       |                                            | ④ 0.7366                                                                                                                                                                                                                                                                                                                                                                                                                     |                                                                                                              |                        |
| 2023/Cao, Bing,<br>et al. [43]                                     | BraTS 2020/494 (369<br>training, 125 testing)/MRI<br>(T1, FLAIR, T1c and T2)<br>ISLES 2015/(28 training,<br>17 testing)/MRI (T1,<br>FLAIR, DWI and T2)<br>CBMFM/69 (62 training,<br>7 testing)/T1 MRI, CT | Transfer target-specific<br>information to generator<br>with multi-scale feature<br>constraints and use an<br>auto-encoder-based<br>discriminator                                                               | GAN, auto-en-<br>coder | -(BraTS 2020 and ISLES<br>2015) resizing<br>-(CBMFM) zero-padding,<br>center-cropping | ① synthesized image<br>quality measurement | PSNR; SSIM; Feature similar-<br>ity index measure (FSIM)<br>① (test on BraTS 2020)<br>toT1<br>26.42; 93.42; 96.86<br>toT1c<br>29.45; 93.16; 96.94<br>toT2<br>26.91; 92.24; 96.34<br>toFLAIR<br>26.43; 90.44; 96.07<br>(test on ISLES 2015)<br>toT1<br>24.24; 87.70; 94.47<br>toDWI<br>29.09; 91.35; 95.78<br>toT2<br>24.94; 85.33; 93.25<br>toFLAIR<br>29.08; 89.47; 95.07<br>(test on CBMFM)<br>toCT<br>23.30; 83.34; 94.56 | <a href="https://github.com/bcaosudo/bcaosudo-AE-GAN">https://<br/>github.com/<br/>bcaosudo/<br/>AE-GAN.</a> | X                      |
|                                                                    |                                                                                                                                                                                                           |                                                                                                                                                                                                                 |                        |                                                                                       |                                            |                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                                              |                        |
|                                                                    |                                                                                                                                                                                                           |                                                                                                                                                                                                                 |                        |                                                                                       |                                            |                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                                              |                        |
|                                                                    |                                                                                                                                                                                                           |                                                                                                                                                                                                                 |                        |                                                                                       |                                            |                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                                              |                        |
|                                                                    |                                                                                                                                                                                                           |                                                                                                                                                                                                                 |                        |                                                                                       |                                            |                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                                              |                        |
|                                                                    |                                                                                                                                                                                                           |                                                                                                                                                                                                                 |                        |                                                                                       |                                            |                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                                              |                        |
|                                                                    |                                                                                                                                                                                                           |                                                                                                                                                                                                                 |                        |                                                                                       |                                            |                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                                              |                        |
|                                                                    |                                                                                                                                                                                                           |                                                                                                                                                                                                                 |                        |                                                                                       |                                            |                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                                              |                        |
|                                                                    |                                                                                                                                                                                                           |                                                                                                                                                                                                                 |                        |                                                                                       |                                            |                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                                              |                        |
|                                                                    |                                                                                                                                                                                                           |                                                                                                                                                                                                                 |                        |                                                                                       |                                            |                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                                              |                        |
|                                                                    |                                                                                                                                                                                                           |                                                                                                                                                                                                                 |                        |                                                                                       |                                            |                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                                              |                        |
|                                                                    |                                                                                                                                                                                                           |                                                                                                                                                                                                                 |                        |                                                                                       |                                            |                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                                              |                        |
|                                                                    |                                                                                                                                                                                                           |                                                                                                                                                                                                                 |                        |                                                                                       |                                            |                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                                              |                        |

Table 4  
(Continued)

| Image synthesis methods              |                                                                                                                                                                                   |                                                                                                                                                                                                                                      |                       |                                                                                                                                  |                                                                             |                                                                                                                                                                                                                                                 |      |                        |
|--------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|----------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|------------------------|
| Year/<br>author/<br>citation         | Dataset/no. of subjects/<br>modalities                                                                                                                                            | Main context                                                                                                                                                                                                                         | Architecture          | Preprocessing                                                                                                                    | Task                                                                        | Metrics/performance                                                                                                                                                                                                                             | Code | External<br>validation |
| 2023/Wang,<br>Yulin, et al. [44]     | IXI/578 (405 training, 58<br>validation, 115 testing)/T1<br>MRI, T2 MRI, (PD)<br>BraTS 2021/1251 (500<br>training, 72 validation,<br>142 testing)/MRI (T1,<br>FLAIR, T1CE and T2) | Utilize a hybrid gener-<br>ator and a CNN-based<br>discriminator                                                                                                                                                                     | CNN, trans-<br>former | -(IXI)<br>spatial registration, cropping<br>-(BraTS)<br>Cropping<br>-(both)<br>Image blurring, down-sam-<br>pling, interpolation | ① random synthesis<br>(SYN)<br>② super-resolution<br>(SR)<br>③ segmentation | PSNR; SSIM                                                                                                                                                                                                                                      | N/A  | X                      |
|                                      |                                                                                                                                                                                   |                                                                                                                                                                                                                                      |                       |                                                                                                                                  |                                                                             | ① (test on BraTS 2021)<br>to T1CE<br>31.07; 0.949<br>(test on IXI)<br>to PD<br>29.95; 0.945<br>② (test on BraTS 2021)<br>to T1CE<br>35.70; 0.983<br>(test on IXI)<br>to PD<br>33.41; 0.975<br>DSC<br>③ (test on BraTS 2021)<br>to T1CE<br>0.917 |      |                        |
| 2023/Liu, Jiang,<br>et al. [45]      | IXI/577 (521 training, 28<br>validation, 28 testing)/T1<br>MRI, T2 MRI, PD<br>BraTS 2021/1251 (1123<br>training, 63 validation, 63<br>testing)/MRI (T1, FLAIR,<br>DWI and T2)     | Treat missing modality<br>imputation as a se-<br>quence-to-sequence task,<br>utilizing a transformer<br>architecture that captures<br>multi-contrast and multi-<br>scale information                                                 | Transformer           | -(IXI)<br>Co-registration<br>-(BraTS)<br>Exclusion, cropping<br>-(both)<br>Normalization                                         | ① synthesized image<br>quality measurement                                  | PSNR; SSIM; LPIPS<br>① (test on BraTS 2021)<br>missing one modal<br>27.87; 0.932; 0.104<br>(test on IXI)<br>Missing one modal<br>36.31; 0.961; 0.080                                                                                            | N/A  | X                      |
| 2023/Wu,<br>Jianghao, et al.<br>[46] | BraTS 2020/369 (258<br>training, 37 validation, 74<br>testing)/MRI (T1, FLAIR,<br>DWI and T2)                                                                                     | Train and infer with<br>cascaded dual-task<br>architecture; use coarse<br>segmentation to regular-<br>ize the synthesis process<br>and apply a tumor-in-<br>formed loss function in<br>discriminator to enhance<br>synthesis quality | U-net                 | Cropping, normalization                                                                                                          | ① synthesized image<br>quality measurement<br>② segmentation (WT)           | Global SSIM; Local SSIM;<br>Global PSNR; Local PSNR<br>① missing FLAIR<br>0.75; 0.52; 22.99; 20.01<br>DSC<br>② missing FLAIR<br>86.09                                                                                                           | N/A  | X                      |
| 2022/Jin,<br>Leiming, et al.<br>[47] | ADNI-1, ADNI2, AD-<br>NI-GO/360/MRI, PET                                                                                                                                          | Use a pretrained classi-<br>fication network to guide a<br>GAN for image synthesis                                                                                                                                                   | GAN, U-net,<br>Resnet | Registration, skull stripping,<br>segmentation, removal, resiz-<br>ing, normalization                                            |                                                                             |                                                                                                                                                                                                                                                 | N/A  | X                      |

Table 4  
(Continued)

| Image synthesis methods                                 |                                                                                                                                            |                                                                                                                                                                                                                                                                  |                          |                                                                                                                         |                                                                                                             |                                                                                                                                                                                                                                                |                                                                                                  |                        |
|---------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------|-------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|------------------------|
| Year/<br>author/<br>citation                            | Dataset/no. of subjects/<br>modalities                                                                                                     | Main context                                                                                                                                                                                                                                                     | Architecture             | Preprocessing                                                                                                           | Task                                                                                                        | Metrics/performance                                                                                                                                                                                                                            | Code                                                                                             | External<br>validation |
| 2022/Sun,<br>Yuqing, Yong<br>Liu, and Bing<br>Liu. [48] | ADNI/363 (223 sCN, 9<br>pCN, 131 MCI)/T1 MRI,<br>PET, clinical data                                                                        | Reconstruct missing tau<br>PET images from paired<br>T1 MRI data                                                                                                                                                                                                 | GAN                      | -(MRI)<br>Segmentation, normalization<br>-(PET)<br>Co-registration, normalization<br>-(all)<br>Reslicing, normalization | Prediction<br>⊙ CN vs. MCI                                                                                  | AUC; ACC; SEN; SPE; FIS<br>⊙ missing PET<br>0.715; 0.713; 0.529; 0.848;<br>0.610                                                                                                                                                               | N/A                                                                                              | X                      |
| 2022/Zhang, Jin,<br>et al. [49]                         | ADNI/873 (677 training,<br>31 validation, 165 test-<br>ing)/T1 MRI, FDG-PET                                                                | Use 3D U-Net and<br>generator with a 3D<br>gradient profile loss and<br>SSIM loss                                                                                                                                                                                | GAN, U-net               | Registration, alignment, re-<br>moval, resizing, normalization                                                          | ⊙ synthesized image<br>quality measurement                                                                  | Mean absolute error (MAE);<br>PSNR; SSIM<br>⊙ missing PET, not splited by<br>subjectID<br>0.0318; 26.92; 0.7294<br>Missing PET, splited by<br>subjectID<br>0.0396; 25.08; 0.6646                                                               | N/A                                                                                              | X                      |
| 2022/Liu, Yunbi,<br>et al. [50]                         | CLAS/76/T1 MRI<br>ADNI-1, ADNI-2/1145/<br>T1 MRI, FDG-PET<br>AIBL/235/T1 MRI,<br>Flutemetamol Pittsburgh<br>compound B (Flute PIB)-<br>PET | Combine two subnet-<br>works, image synthesis<br>and representation<br>learning, sharing the<br>same imaging features;<br>use transfer learning<br>to handle limited data<br>problem                                                                             | GAN                      | -(MRI)<br>Skull-stripping, correction,<br>normalization<br>-(PET)<br>Alignment, normalization                           | Classification<br>⊙ progressive subjec-<br>tive cognitive decline<br>(SCD) vs stable SCD<br>⊙ pMCI vs. sMCI | AUC; Balanced accuracy<br>(BAC); SPE; SEN; FIS<br>⊙ missing PET (test on<br>CLAS)<br>0.747; 0.721; 0.692; 0.750;<br>0.621<br>AUC; ACC<br>⊙ missing PET (test on ADNI)<br>0.838; 0.780                                                          | <a href="https://github.com/Candyeeee/JSRL">https://<br/>github.com/<br/>Candyeeee/<br/>JSRL</a> | ✓                      |
| 2022/Huang, Pu,<br>et al. [51]                          | BraTS 2019/335 (80%<br>training, 10% validation,<br>10% testing)/MRI (T1,<br>FLAIR, T1c and T2)                                            | Uses different encoders<br>to separately process<br>each type of medical<br>image and extract im-<br>portant features and then<br>use group convolution to<br>combine these features<br>from different image<br>types into a single, uni-<br>fied representation | GAN, encoder-<br>decoder | Alignment, normalization                                                                                                | ⊙ synthesized image<br>quality measurement                                                                  | PSNR; SSIM; Normalized<br>mean squared error (NMSE)<br>⊙ T1toT2<br>(whole image)<br>27.38; 0.953; 0.030<br>(tumor region)<br>18.08; 0.686; 0.007<br>T1toFLAIR<br>(whole image)<br>28.41; 0.962; 0.014<br>(tumor region)<br>17.81; 0.607; 0.007 | N/A                                                                                              | X                      |
| 2022/Kaplan,<br>Sydney, et al.<br>[52]                  | eLABLE/127/T1 MRI, T2<br>MRI, rs-fMRI<br>ECHO/10/T1 MRI, T2<br>MRI, rs-fMRI                                                                | Train 3D-GAN on full<br>volumetric data follow-<br>ing cycleGAN procedure                                                                                                                                                                                        | GAN                      | Distortion correction, denois-<br>ing, normalization, registra-<br>tion, filtering                                      | ⊙ synthesized image<br>quality measurement                                                                  | MAE; Mean structural similar-<br>ity index measure (MSSIM);<br>DSC<br>⊙ T1toT2<br>0.056; 0.79; 0.82                                                                                                                                            | N/A                                                                                              | ✓                      |

Table 4  
(Continued)

| Image synthesis methods                                                             |                                                                                                                                     |                                                                                                                                                                                                                                |                                         |                                                                                                                               |                                                                                                |                                                                                                                                                                                                                                                     |      |                        |
|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------|-------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|------------------------|
| Year/<br>author/<br>citation                                                        | Dataset/no. of subjects/<br>modalities                                                                                              | Main context                                                                                                                                                                                                                   | Architecture                            | Preprocessing                                                                                                                 | Task                                                                                           | Metrics/performance                                                                                                                                                                                                                                 | Code | External<br>validation |
| 2021/Gao,<br>Xingyu, et al.<br>[53]                                                 | ADNI-1/821 (196 AD,<br>168 pMCI, 230 sMCI),227<br>CN)/T1 MRI, FDG-PET                                                               | Combine pyramid<br>convolution and attention<br>module with a task-in-<br>duced discriminator                                                                                                                                  | GAN, U-net                              | Skull-stripping, registration,<br>alignment, cropping, down-<br>sampling                                                      | ① synthesized image<br>quality measurement<br>classification<br>② AD vs. CN<br>③ pMCI vs. sMCI | SSIM; PSNR; Mean squared<br>error (MSE); Maximum mean<br>discrepancy (MMD)<br>① missing PET<br>0.915; 29.0; 184; 0.107<br>ACC; SEN; SPE; AUC; FIS<br>② missing PET<br>92.0; 89.1; 94.0; 95.6; 90.5<br>③ missing PET<br>75.3; 77.3; 74.1; 78.6; 69.9 | N/A  | X                      |
|                                                                                     | ADNI-2/534 (156 AD,<br>66 pMCI, 112 sMCI, 200<br>CN)/T1 MRI, FDG-PET                                                                |                                                                                                                                                                                                                                |                                         |                                                                                                                               |                                                                                                |                                                                                                                                                                                                                                                     |      |                        |
| 2021/Yan, Kun,<br>et al [54]                                                        | BraTS2015/54 (42<br>trainings, 6 validation, 6<br>testing)/MRI (T1, FLAIR,<br>T1c and T2)                                           | Use a semi-supervised<br>cross-modal MRI syn-<br>thesis network based on<br>cycleGAN adopting the<br>coarse-to-fine learning<br>strategy                                                                                       | GAN, denoising<br>autoencoders<br>[101] | Normalization, cropping                                                                                                       | ① synthesized image<br>quality measurement                                                     | MSE; SSIM; PSNR<br>① (20%paired)<br>T1toT2<br>0.0054; 0.9047; 23.85<br>T2toT1<br>0.0049; 0.9207; 23.79                                                                                                                                              | N/A  | X                      |
|                                                                                     |                                                                                                                                     |                                                                                                                                                                                                                                |                                         |                                                                                                                               |                                                                                                |                                                                                                                                                                                                                                                     |      |                        |
| 2021/Islam,<br>Mobarakol,<br>Navodini<br>Wijethilake, and<br>Hongliang Ren.<br>[55] | BraTS 2017/331 (285<br>trainings, 46 testing)/MRI<br>(T1, FLAIR, T1c and T2)<br>TCGA/202/MRI (T1,<br>FLAIR and T2), genomic<br>data | Use a network that trans-<br>lates multi-modal MRI<br>inputs into a single MRI<br>modality, employing<br>a fully convolutional<br>network as the generator<br>and an encoder-decod-<br>er architecture as the<br>discriminator | GAN, FCN                                | Resampling, registration                                                                                                      | ① synthesized image<br>quality measurement<br>segmentation<br>② WT<br>③ TC<br>④ ET             | PSNR<br>① T1,FLAIRtoT2<br>24.9387<br>DSC<br>② 0.8776<br>③ 0.7615<br>④ 0.7243                                                                                                                                                                        | N/A  | ✓                      |
|                                                                                     |                                                                                                                                     |                                                                                                                                                                                                                                |                                         |                                                                                                                               |                                                                                                |                                                                                                                                                                                                                                                     |      |                        |
| 2021/Akbar,<br>Muhammad<br>Usman, Vittorio<br>Murino, and<br>Diego Sona. [56]       | Synapse/30 (24 training, 2<br>validation, 4 testing)/CT<br>CHAOS/20 (16 training,<br>2 validation, 2 testing)/T1<br>MRI, T2 MRI)    | Enhance the dataset by<br>synthesizing one imaging<br>modality from another<br>modality sourced from a<br>separate dataset                                                                                                     | cycle-GAN                               | Re-sampling of CT, label<br>balancing, normalization                                                                          | Segmentation<br>① Liver<br>② Spleen<br>③ Right kidney<br>④ Left kidney                         | DSC<br>① CT 0.9519; MRI 0.9347<br>② CT 0.9215; MRI 0.8463<br>③ CT 0.8695; MRI 0.8948<br>④ CT 0.8891; MRI 0.8330                                                                                                                                     | N/A  | X                      |
|                                                                                     |                                                                                                                                     |                                                                                                                                                                                                                                |                                         |                                                                                                                               |                                                                                                |                                                                                                                                                                                                                                                     |      |                        |
| 2021/Lin,<br>Wanyun, et al.<br>[57]                                                 | ADNI-1,2,3,GO/1086<br>(70% training, 20%val-<br>idation, 10% testing)/<br>MRI,pPET                                                  | Use reversible structure<br>and improve the genera-<br>tor part in GAN                                                                                                                                                         | 3D GAN                                  | -(MRI)<br>Correction, resampling,<br>cropping<br>-(PET)<br>Registration, cropping,<br>normalization, averaging,<br>resampling | Classification<br>① AD vs. CN<br>② pMCI vs. sMCI                                               | ACC; SEN; SPE; AUC<br>① missing PET<br>89.05; 90.48; 87.50; 87.92<br>Missing MRI<br>88.64; 91.60; 85.71; 87.36<br>② missing PET<br>71.23; 74.36; 67.65; 73.66<br>Missing MRI<br>71.18; 69.07; 73.97; 70.80                                          | N/A  | X                      |
|                                                                                     |                                                                                                                                     |                                                                                                                                                                                                                                |                                         |                                                                                                                               |                                                                                                |                                                                                                                                                                                                                                                     |      |                        |

Table 4  
(Continued)

| Image synthesis methods               |                                                                                                                                                                                                                                                                                                                                |                                                                                                                                               |                          |                                                                   |                                                                                                |                                                                                                                                                                                                                                                                                                                 |      |                        |
|---------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------|--------------------------|-------------------------------------------------------------------|------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|------------------------|
| Year/<br>author/<br>citation          | Dataset/no. of subjects/<br>modalities                                                                                                                                                                                                                                                                                         | Main context                                                                                                                                  | Architecture             | Preprocessing                                                     | Task                                                                                           | Metrics/performance                                                                                                                                                                                                                                                                                             | Code | External<br>validation |
| 2020/Pan,<br>Yongsheng, et al. [58]   | ADNI-1,2; the Australian<br>Imaging/2355/MRI, PET                                                                                                                                                                                                                                                                              | Invent a hybrid loss<br>function containing three<br>losses: voxel-wise-con-<br>sistent loss, cycle-consis-<br>tent loss, adversarial loss    | GAN                      | N/A                                                               | ① synthesized image<br>quality measurement<br>classification<br>② AD vs. CN<br>③ pMCI vs. sMCI | PSNR; SSIM; MAE<br>① toMRI<br>26.07; 0.6683; 0.1070<br>toPET<br>30.24; 0.6945; 0.0757<br>ACC; AUC; SEN; SPE; F1S;<br>Matthews correlation coeffi-<br>cient (MCC)<br>② 93.58; 96.95; 91.52; 95.22;<br>92.64; 86.97<br>③ 77.44; 82.51; 79.07; 77.22;<br>45.64; 40.06                                              | N/A  | X                      |
| 2020/Liu, Yunbi,<br>et al. [59]       | ADNI, CLAS/1055 (863<br>training (ADNI), 79 test-<br>ing (ADNI), 113 testing<br>(CLAS))/MRI, PET                                                                                                                                                                                                                               | Combine image synthesis<br>network with a represen-<br>tation learning network;<br>use transfer learning<br>to handle limited data<br>problem | 3D-GAN                   | Skull-stripping, intensity cor-<br>rection, spatial normalization | Classification<br>① stable SCD vs.<br>progressive SCD                                          | AUC; ACC; SPE; SEN; F1S<br>② 0.713; 0.655; 0.616; 0.725;<br>0.598                                                                                                                                                                                                                                               | N/A  | ✓                      |
| 2020/Dar, Sal-<br>man UH, et al. [60] | MIDAS/40 (25 training, 5<br>validation, 10 testing)/T1<br>MRI, T2 MRI<br>IXI/40 (25 training, 5<br>validation, 10 testing)/T1<br>MRI, T2 MRI, PD<br>BraTS2015/40 (25<br>training, 5 validation, 10<br>testing)/T1 MRI, T2 MRI,<br>FLAIR<br>Multi-Coil MRI/10 (7<br>training, 1 validation, 2<br>testing)/T1 MRI, T2 MRI,<br>PD | Use two branches to<br>jointly reconstruct and<br>synthesize the target<br>contrast                                                           | Conditional<br>GAN [102] | Discarding, registration,<br>upsampling                           | ① synthesized image<br>quality measurement                                                     | PSNR; SSIM<br>① (test on MIDAS)<br>T1toT2<br>37.35; 97.96<br>T2toT1<br>32.17; 94.53<br>(test on IXI)<br>T1toT2<br>36.18; 97.70<br>T2toT1<br>34.56; 97.91<br>(test on BraTS2015)<br>T1toT2<br>36.15; 98.77<br>T2toT1<br>36.46; 98.86<br>(test on Multi-Coil)<br>T1toT2<br>35.51; 97.61<br>T2toT1<br>36.40; 97.74 | N/A  | X                      |

Table 4  
(Continued)

| Image synthesis methods                                               |                                                                                                                                                    |                                                                                                                                                                  |                   |                                                    |                                                                    |                                                                                                                                                  |      |                        |
|-----------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|----------------------------------------------------|--------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------|------|------------------------|
| Year/<br>author/<br>citation                                          | Dataset/no. of subjects/<br>modalities                                                                                                             | Main context                                                                                                                                                     | Architecture      | Preprocessing                                      | Task                                                               | Metrics/performance                                                                                                                              | Code | External<br>validation |
| 2020/Cao, Bing,<br>et al. [61]                                        | ADNI/16/T1 MRI, CT<br>BraTS/352 (80% training,<br>20% testing)/MRI (T1,<br>FLAIR, T1c and T2)                                                      | Use auto-encoder<br>network for self-supervised<br>learning, and use<br>a collaborative learning<br>framework to utilize information<br>from multiple modalities | GAN, auto-encoder | N/A                                                | ① synthesized image<br>quality measurement                         | SSIM; FSIM<br>① missing T1<br>0.9317; 0.9657<br>Missing T1c<br>0.9061; 0.9573<br>Missing FLAIR<br>0.8948; 0.9602<br>Missing T2<br>0.9157; 0.9592 | N/A  | X                      |
| 2019/Sharma,<br>Anmol, and<br>Ghassan Hamarneh. [62]                  | ISLES2015/22 (training)/<br>MRI<br>(T1, T2, DWI, FLAIR)<br>BraTS2018/285 (260<br>training, 10 validation, 15<br>testing)/MRI (T1,T2,T1C,<br>FLAIR) | Use multi-input<br>multi-output synthesizer<br>using implicit conditioning<br>and trained using<br>curriculum learning                                           | GAN, U-net        | Mean normalization, cropping,<br>resizing          | ① synthesized image<br>quality measurement                         | MSE; PSNR; SSIM<br>① (mean)<br>Missing T1<br>0.0052; 26.6057; 0.9276<br>Missing T2<br>0.0049; 26.1233; 0.9078                                    | N/A  | X                      |
| 2018/Pan,<br>Yongsheng, et<br>al. [63]                                | ADNI-1,2/821/MRI, PET                                                                                                                              | Impute missing modalities<br>with cycle-consistent<br>GAN to learn bi-directional<br>mapping between different<br>modalities                                     | 3D-cGAN           | Alignment, skull stripping,<br>correction, removal | Classification<br>① AD vs. healthy control (HC)<br>② pMCI vs. sMCI | ACC; SEN; SPE; FIS; MCC;<br>AUC<br>① 92.50; 89.94; 94.53; 91.37;<br>84.78; 95.89<br>② 79.08; 55.26; 82.85; 40.86;<br>30.13; 75.84                | N/A  | X                      |
| 2018/Tang,<br>Zhenyu,<br>Pew-Thian Yap,<br>and Dinggang<br>Shen. [64] | BraTS2015<br>100 (50 images out of<br>1000 for testing)/MRI<br>(T1,T2,T1C, FLAIR)<br>LPBA40/T1 MRI                                                 | Use synthesizers to generate<br>multimodal normal<br>atlases from standard<br>single modality normal<br>atlases                                                  | cycleGAN          | Affine transformation, histogram<br>matching       | ① segmentation                                                     | DSC (average)<br>① gray matter<br>0.701<br>White matter<br>0.748<br>CSF<br>0.593                                                                 | N/A  | X                      |

Table 5  
Knowledge transfer methods dealing with missing imaging modality

| Knowledge transfer methods            |                                                                                                                                                          |                                                                                                                                                                                                                                |                                     |                                                                                                                                                                                                                                                                                          |                                                                                                                                             |                                                                                                                                                                                                                                                                                                                                  |
|---------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Year/<br>author/<br>citation          | Dataset/no. of<br>subjects/modal-<br>ities                                                                                                               | Main context                                                                                                                                                                                                                   | Architec-<br>ture                   | Preprocessing                                                                                                                                                                                                                                                                            | Task                                                                                                                                        | Metrics/performance                                                                                                                                                                                                                                                                                                              |
| 2025/Liu,<br>et al. [65]              | BraTS2020, 2018,<br>2015/927 (660<br>training, 91 valida-<br>tion, 177 testing)/<br>MRI (T1, FLAIR,<br>T1c and T2)                                       | Propose an enhanced model<br>with two key stages: a pretrain-<br>ing phase synthesizing diverse<br>MRI data by decoupling anatomi-<br>my and tumor components, and<br>a post-training phase applying<br>knowledge distillation | 3D encoder-<br>decoder, 3D<br>U-net | Background removal,<br>normalization, crop-<br>ping, augmentations                                                                                                                                                                                                                       | Segmentation<br>① whole tumor<br>(WT)<br>② tumor core (TC)<br>③ enhanced tumor<br>(ET)                                                      | Dice similarity coefficient (DSC)<br>① missing T1 (complete data)<br>91.15 (91.29)<br>② missing T1 and T2 (complete data)<br>86.21 (85.91)<br>③ missing T1 (complete data)<br>82.41 (82.09)<br><br>Code<br><a href="https://github.com/ZhongAo-bo/Asymmetry-BTS">https://github.com/<br/>ZhongAo-<br/>bo/Asym-<br/>metry-BTS</a> |
| 2023/Xing,<br>Xiaohan,<br>et al. [66] | TCGA-GBM,<br>TCGA-LGG<br>project/737 (182<br>grade II, 205 grade<br>III, 350 grade<br>IV; 80%training,<br>20%testing)/ROI<br>images, genomic<br>features | Multi-modal teacher network<br>trained for adaptive teaching of<br>unimodal student network                                                                                                                                    | ResNet                              | Augmentation (random<br>cropping, color jitter-<br>ing, flipping)                                                                                                                                                                                                                        | ① pathological<br>glioma grading                                                                                                            | Area under the curve(AUC); Average<br>precision (AP); Accuracy (ACC);<br>Kappa score<br>① missing<br>92.42, 86.34, 76.47, 64.34<br><br><a href="https://github.com/CUHK- AIMGroup/ MultiMod- al-learning">https://<br/>github.com/<br/>CUHK-<br/>AIMGroup/<br/>MultiMod-<br/>al-learning</a> .                                   |
| 2023/Chen,<br>Yuan,<br>et al. [67]    | ADNI-1, ADNI-2,<br>ADNI-GO/1248<br>(347 AD, 417 CN,<br>484 MCI; 80%<br>training, 20%<br>testing)/T1 sMRI,<br>FDG PET                                     | Use knowledge distillation<br>model to impute representation<br>of the missing modality                                                                                                                                        | CNN                                 | -(sMRI)<br>Reorientation, normal-<br>ization, registration,<br>skull stripping, cere-<br>bellum removal<br>-(PET)<br>Smoothing, coreg-<br>istration, averaging,<br>computing<br>AC-PC orient baseline,<br>standardization, nor-<br>malization<br>-(both)<br>Resizing, normaliza-<br>tion | Classification<br>① Alzheimer's<br>disease (AD) vs.<br>Cognitively Nor-<br>mal (CN)<br>② progressive<br>MCI (pMCI) vs.<br>stable MCI (sMCI) | AUC; AP; Sensitivity (SEN); Spec-<br>ificity (SPE); Matthews correlation<br>coefficient (MCC)<br>① missing<br>96.85; 90.23; 91.73; 93.69; 84.21<br>② missing<br>83.81; 77.26; 72.97; 79.02; 51.27<br><br>N/A                                                                                                                     |
|                                       |                                                                                                                                                          |                                                                                                                                                                                                                                |                                     |                                                                                                                                                                                                                                                                                          |                                                                                                                                             | X                                                                                                                                                                                                                                                                                                                                |

Table 5  
(Continued)

| Knowledge transfer methods               |                                                                                                                                                                                                  |                                                                                                                                                                               |                           |                                                                                                                                                                             |                                                                                        |                                                                                                                                                                                                                       |                                                                                                                          |                        |
|------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|------------------------|
| Year/<br>author/<br>citation             | Dataset/no. of<br>subjects/modal-<br>ities                                                                                                                                                       | Main context                                                                                                                                                                  | Architec-<br>ture         | Preprocessing                                                                                                                                                               | Task                                                                                   | Metrics/performance                                                                                                                                                                                                   | Code                                                                                                                     | External<br>validation |
| 2023/Dolci,<br>G., et al.<br>[68]        | ADNI/1581<br>(644CN,332AD,<br>316MCI<br>non-converters<br>(MCIInc), 289MCI<br>converters (MCIc);<br>80% training, 20%<br>testing)/sMRI,<br>fMRI, Single nu-<br>cleotide polymor-<br>phisms (SNP) | Recover missing modalities<br>using pre-trained generators in<br>the latent space                                                                                             | cycleGAN                  | -(sMRI) normalization,<br>smoothing                                                                                                                                         | Classification<br>① AD vs. CN                                                          | ACC; Precision (PRE); Recall<br>(REC)                                                                                                                                                                                 | N/A                                                                                                                      | X                      |
|                                          |                                                                                                                                                                                                  |                                                                                                                                                                               |                           | -(fMRI)<br>rigid body motion<br>correction, slice-timing<br>correction, warping,<br>resampling<br>-(genomics data) gen-<br>otyped, Pre-imputation<br>QC, imputation         | ② MCIInc vs.<br>MCIc                                                                   | ① missing<br>0.938; 0.905; 0.884<br>② missing<br>0.716; 0.622; 0.730                                                                                                                                                  |                                                                                                                          |                        |
| 2022/Dolci,<br>Giorgio, et<br>al. [69]   | ADNI/788 (80%<br>training and valida-<br>tion, 20% testing)/<br>sMRI, fMRI, SNP                                                                                                                  | Use two cycleGAN to perform<br>knowledge transfer from avail-<br>able modalities                                                                                              | cycleGAN                  | -(fMRI)<br>correction, warping,<br>resampling, smoothing<br>-(sMRI)<br>segmentation, smooth-<br>ing<br>-(both MRI)<br>removal<br>-(SNP)<br>pre-imputation QC,<br>imputation | Classification<br>① AD vs. CN                                                          | ACC; REC; PRE<br>① 87% at least one missing modal<br>0.935; 0.894; 0.894                                                                                                                                              | N/A                                                                                                                      | X                      |
| 2022/Jeong,<br>Seung-wan,<br>et al. [70] | BraTS2017/285/<br>MRI (T1, FLAIR,<br>T1c and T2)                                                                                                                                                 | Use adversarial learning to gen-<br>erate the missing latent features<br>and create a shared represen-<br>tation with an attention-based<br>fusion block                      | ResNet                    | Skull stripping,<br>co-registration, resiz-<br>ing, normalization,<br>augmentation (flipping,<br>rotation)                                                                  | ① tumor classifi-<br>cation                                                            | ACC; AUC; SEN; SPE<br>① missingT1 (complete data)<br>88.75 (90.91);<br>missingT1 (complete data)<br>95.58 (96.34);<br>missingT2,FLAIR (complete data)<br>92.12 (92.69);<br>missingT1 (complete data)<br>83.53 (84.90) | N/A                                                                                                                      | X                      |
| 2022/Yang,<br>Qiushi, et<br>al. [71]     | BraTS2018/285<br>(195 training, 90<br>validation)/MRI<br>(T1, FLAIR, T1c<br>and T2)                                                                                                              | Use a method where two<br>models learn together by<br>sharing detailed tumor-region<br>information, guided by how<br>similar different regions are, to<br>improve performance | U <sup>2</sup> -net [105] | Normalization, random<br>cropping, augmenta-<br>tion (random flipping,<br>random rotation, ran-<br>dom intensity change)                                                    | Segmentation<br>① whole tumor<br>(WT)<br>② tumor core (TC)<br>③ enhanced tumor<br>(ET) | DSC<br>① missing T1 (complete data)<br>88.8 (88.8)<br>② missing T2 (complete data)<br>80.9 (80.1)<br>③ missing T1,FLAIR (complete data)<br>68.7 (68.4)                                                                | <a href="https://github.com/CityU-AIM-Group/D2Net">https://<br/>github.<br/>com/CityU-<br/>AIM-<br/>Group/<br/>D2Net</a> | X                      |

Table 5  
(Continued)

| Knowledge transfer methods                      |                                                                                                    |                                                                                                                                                                                 |                   |                                                             |                                      |                                                                                                                                                                                                                                   |
|-------------------------------------------------|----------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|-------------------------------------------------------------|--------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Year/<br>author/<br>citation                    | Dataset/no. of<br>subjects/modal-<br>ities                                                         | Main context                                                                                                                                                                    | Architec-<br>ture | Preprocessing                                               | Task                                 | External<br>validation                                                                                                                                                                                                            |
| 2022/Li,<br>Haoran, et<br>al. [72]              | BraTS2020/369<br>(220 training, 74<br>validation, 75<br>testing)/MRI (T1,<br>FLAIR, T1c and<br>T2) | Use a Bernoulli sampling<br>process to create training inputs<br>for the student network; use a<br>supervised knowledge transfer<br>loss between teacher and stu-<br>dent model | 3D U-net          | Co-registration, inter-<br>polation, skull-strip-<br>ping   | Segmentation<br>① WT<br>② TC<br>③ ET | DSC<br>① missing T1 (complete data)<br>89.7 (90.2)<br>② missing T1 (complete data)<br>87.2 (87.0)<br>③ missing FLAIR (complete data)<br>81.2 (82.3)                                                                               |
| 2021/<br>Vadachino,<br>Saverio, et<br>al. [73]  | BraTS2019/460<br>(335 training, 125<br>validation)/MRI<br>(T1, FLAIR, T1c<br>and T2)               | Use a hierarchical discriminator<br>which distills the latent infor-<br>mation to overcome the large<br>domain shift problem                                                    | 3D U-net          | N/A                                                         | Segmentation<br>① WT<br>② TC<br>③ ET | DSC<br>① missing T1c<br>87.5<br>② missing T1c<br>66.7<br>③ missing T1c<br>39.8                                                                                                                                                    |
| 2021/Wang,<br>Yixin, et al.<br>[74]             | BraTS2018/285/<br>MRI (T1, FLAIR,<br>T1c and T2)                                                   | Use co-training approach<br>between network composed of<br>multimodal and unimodal paths                                                                                        | U-net             | Co-registration, inter-<br>polation                         | Segmentation<br>① WT<br>② TC<br>③ ET | DSC; Hausdorff distance (HD95)<br>① missing T2 (complete data)<br>88.96 (89.22); 6.93 (6.71)<br>② missing T2,FLAIR (complete data)<br>84.59 (85.18); 5.76 (5.94)<br>③ missing T1,T2 (complete data)<br>77.46 (77.06); 4.22 (5.09) |
| 2020/Hu,<br>Minhao, et<br>al. [75]              | BraTS2018/285<br>(2/3 training, 1/3<br>validation)/MRI<br>(T1, FLAIR, T1c<br>and T2)               | Use a generalized knowledge<br>distillation network, by con-<br>straining the latent representa-<br>tion of a mono-modal 'student'<br>to be similar to multi-modal<br>'teacher' | U-net             | Cropping, random<br>flipping, normalization,<br>up-sampling | Segmentation<br>① WT<br>② TC<br>③ ET | DSC (only T1c)<br>① 76.98<br>② 81.45<br>③ 71.67                                                                                                                                                                                   |
| 2019/Shen,<br>Yan, and<br>Mingchen<br>Gao. [76] | BraTS2018/361<br>(80% training, 20%<br>validation)/MRI<br>(T1, FLAIR, T1c<br>and T2)               | Use a loss function to make the<br>internal features from missing<br>modalities resemble those from<br>complete ones                                                            | U-net             | Normalization, crop-<br>ping                                | Segmentation<br>① WT<br>② TC<br>③ ET | DSC<br>① missing T2 (complete data)<br>0.893 (0.894)<br>② missing T1 (complete data)<br>0.778 (0.790)<br>③ missing T2 (complete data)<br>0.643 (0.653)                                                                            |

Table 6  
Latent feature space-based methods dealing with missing imaging modality

| Latent feature space-based methods  |                                                                                                                             |                                                                                                                                                                                                                    |                      |                                                                                                                                                 |                                                                                                                                           |                                                                                                                                                                                                                                                                                                                                                                    |                                                                                                  |                        |
|-------------------------------------|-----------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------|-------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|------------------------|
| Year/author/<br>citation            | Dataset/no. of<br>subjects/modalities                                                                                       | Main context                                                                                                                                                                                                       | Architecture         | Preprocessing                                                                                                                                   | Task                                                                                                                                      | Metrics/performance                                                                                                                                                                                                                                                                                                                                                | Code                                                                                             | External<br>validation |
| 2023/Zhou,<br>Tongxue. [77]         | BraTS 2018/285/MRI<br>(T1, FLAIR, T1c and<br>T2)                                                                            | Integrate information across<br>modalities, leverage latent<br>multimodal correlations, and<br>perform feature extraction                                                                                          | U-net [108]          | Co-registration,<br>interpolation,<br>skull-stripping, manual<br>labeling of ground<br>truth, resizing, bias field<br>correction, normalization | Segmentation<br>① whole tumor<br>(WT)<br>② tumor core (TC)<br>③ enhanced tumor<br>(ET)<br>④ average of three<br>regions                   | Dice similarity coefficient<br>(DSC)<br>① missing T2 (complete data)<br>86.6 (86.5)<br>② missing T2 (complete data)<br>87.1 (87.0)<br>③ missing T2 (complete data)<br>78.8 (78.6)<br>④ missing T2 (complete data)<br>84.1 (84.1)                                                                                                                                   | N/A                                                                                              | X                      |
| 2023/Wang,<br>Tao, et al. [78]      | ADNI-1,2,3, OA-<br>SIS/1530 (1387+ 143)<br>divided into 10<br>subsets (8 training, 1<br>validation, 1 testing)/<br>MRI, PET | Impute missing data from mul-<br>tiple views spanning different<br>modalities and temporal stages,<br>incorporate the adversarial<br>learning into the imputation<br>to make it closer to the real<br>distribution | MinimalRNN<br>[109]  | Correction, skull<br>stripping, registration,<br>segmentation, labeling                                                                         | ① Mild Cognitive<br>Impairment (MCI)<br>conversion predic-<br>tion<br>② imputation assess-<br>ment                                        | Accuracy (ACC); Area under<br>the curve (AUC); Balanced<br>accuracy (BAC)<br>① ADNI1,2<br>0.842; 0.860; 0.830<br>ADNI3<br>0.813; 0.845; 0.821<br>Mean absolute error (MAE);<br>Root mean squared error<br>(RMSE)<br>② ADNI1,2<br>missing MRI<br>0.322; 0.468<br>missing PET<br>0.415; 0.513<br>OASIS<br>missing MRI<br>0.372; 0.519<br>missing PET<br>0.621; 0.738 | <a href="https://github.com/Meiyan88/MCNET">https://<br/>github.com/<br/>Meiyan88/<br/>MCNET</a> | X                      |
| 2023/Gao,<br>Xingyu, et al.<br>[79] | ADNI-1,AD-<br>NI-2/1364/482 incom-<br>plete)/MRI,PET                                                                        | A small network is used to esti-<br>mate important abstract features<br>of the missing data type using<br>regression                                                                                               | Transformer<br>[110] | Skull-stripping, linear<br>registration, downsam-<br>pling                                                                                      | Classification<br>① Alzheimer's<br>disease (AD) vs.<br>cognitively normal<br>(CN)<br>② progressive MCI<br>(pMCI) vs. stable<br>MCI (sMCI) | AUC; ACC; Sensitivity<br>(SEN); Specificity (SPE)<br>① missing PET<br>96.7; 92.4; 88.5; 95.5<br>② missing PET<br>87.2; 77.8; 75.4; 79.6                                                                                                                                                                                                                            | N/A                                                                                              | X                      |

Table 6  
(Continued)

| Latent feature space-based methods    |                                                                                                                    |                                                                                                                                   |                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                                                                                                                                                                             |                                                                                                                                                                                                                                                                                        |                                                                                                  |                        |
|---------------------------------------|--------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------|---------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|------------------------|
| Year/author/<br>citation              | Dataset/no. of<br>subjects/modalities                                                                              | Main context                                                                                                                      | Architecture        | Preprocessing                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | Task                                                                                                                                                                        | Metrics/performance                                                                                                                                                                                                                                                                    | Code                                                                                             | External<br>validation |
| 2023/ Wang,<br>Tao, et al. [80]       | ADNI-1, ADNI-<br>2/3166/T1 MRI,<br>fluorodeoxyglucose<br>(FDG)-PET, SNP<br>PPMI/960/T1 MRI,<br>DTI, SNP            | Genetic data encoded using<br>synthesized vectors to integrate<br>with imaging data representa-<br>tions                          | ResNet [111]        | -(MRI)<br>Anterior commissure<br>posterior correction,<br>image intensity inhom-<br>ogeneity correction, skull<br>stripping, registration,<br>segmentation, ROIs<br>labeling, GM tissue<br>volume computation<br>-(PET)<br>Co-registration<br>-(DTI)<br>b-vector and b-value file<br>generation, eddy cor-<br>rection, skull stripping,<br>fractional anisotropy<br>calculation, alignment,<br>ROIs labeling, mean<br>tissue density calculation<br>-(SNP)<br>Quality control, removal,<br>selection | Classification<br>⓪ AD vs. CN<br>⓪ pMCI vs. sMCI<br>⓪ Parkinson's dis-<br>ease vs. CN<br>⓪ Parkinson's<br>disease vs. scans<br>without evidence for<br>dopaminergic deficit | AUC; ACC; SEN; SPE; Bal-<br>anced classification accuracy<br>(BCA)<br>⓪ missing<br>93.81, 96.98, 93.40, 92.16,<br>93.10<br>⓪ missing<br>76.68, 78.08, 76.40, 77.35,<br>76.88<br>⓪ missing<br>79.96, 81.56, 77.27, 80.62,<br>78.95<br>⓪ missing<br>80.31, 82.51, 88.51, 79.04,<br>83.78 | <a href="https://github.com/Meiyan88/DMAAN">https://<br/>github.com/<br/>Meiyan88/<br/>DMAAN</a> | X                      |
|                                       |                                                                                                                    |                                                                                                                                   |                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                                                                                                                                                                             |                                                                                                                                                                                                                                                                                        |                                                                                                  |                        |
| 2023/Wang,<br>Hu, et al. [81]         | BraTS2018/351 (285<br>training, 66 evalua-<br>tion)/MRI (T1, FLAIR,<br>T1c and T2)                                 | Fuse shared and specific fea-<br>tures of modalities with linear<br>projection                                                    | 3D U-net [112]      | N/A                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | Segmentation<br>⓪ WT<br>⓪ TC<br>⓪ ET                                                                                                                                        | DSC<br>⓪ missing T1 (complete data)<br>90.79 (90.88)<br>⓪ missing T1 (complete data)<br>85.67 (85.75)<br>⓪ missing T1&FLAIR (com-<br>plete data)<br>78.59 (78.47)                                                                                                                      | <a href="https://github.com/billhhh/ShaSpec/">https://github.<br/>com/billhhh/<br/>ShaSpec/</a>  | X                      |
| 2023/Morar,<br>Ulyana, et al.<br>[82] | ADNI/1843 (90%<br>training, 10% testing)/<br>MRI, PET, cerebrospi-<br>nal fluid (CSF), bio-<br>chemical biomarkers | Use a Long short-term memory<br>(LSTM) [113] regressor and a<br>pretrained Neural Network Esti-<br>mator to impute missing values | LSTM                | Augmentation, strati-<br>fying                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | Prediction<br>⓪ Mini-Mental State<br>Examination score                                                                                                                      | Correlation; RMSE; Coeffi-<br>cient of Determination<br>⓪ 90.27; 1.86;81.36                                                                                                                                                                                                            | N/A                                                                                              | X                      |
| 2023/Hou,<br>Wentai, et al.<br>[83]   | TCGA/2102 (60%train-<br>ing, 20%validation,<br>20%testing)/clinical<br>records, genomic data                       | Generate hyperedge using<br>transformer to learn multi-modal<br>dependence                                                        | Hybrid GCN<br>[114] | Normalization                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | Prediction<br>⓪ cancer survival                                                                                                                                             | Concordance index (CI)<br>⓪ KIRC: 0.750<br>LIHC: 0.695<br>ESCA: 0.664<br>LUSC: 0.590<br>LUAD: 0.639<br>UCEC: 0.735                                                                                                                                                                     | <a href="https://github.com/lin-lcx/HGCN">https://github.<br/>com/lin-lcx/<br/>HGCN</a>          | X                      |

Table 6  
(Continued)

| Latent feature space-based methods                                                |                                                                                                               |                                                                                                                                                                    |                     |                                                                                                                 |                                                                            |                                                                                                                                                              | External validation |
|-----------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|-----------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|
| Year/author/<br>citation                                                          | Dataset/no. of<br>subjects/modalities                                                                         | Main context                                                                                                                                                       | Architecture        | Preprocessing                                                                                                   | Task                                                                       | Metrics/performance                                                                                                                                          |                     |
| 2022/Zhang,<br>Yao, et al. [84]                                                   | BraTS 2018/285/MRI<br>(T1, FLAIR, T1c and<br>T2)                                                              | Inter-modal transformer com-<br>bining all the input embeddings,<br>and auxiliary regularizers                                                                     | Transformer,<br>CNN | Co-registration,<br>interpolation,<br>(augmentation)<br>random flipping,<br>cropping,<br>intensity shifts       | Segmentation<br>① WT<br>② TC<br>③ ET                                       | DSC<br>① missing T1 (complete data)<br>88.14 (89.64)<br>② missing FLAIR (complete<br>data)<br>80.39 (85.78)<br>③ missing T1 (complete data)<br>75.67 (77.61) | X                   |
| 2022/<br>Hayat, Nasir,<br>Krzysztof J.<br>Geras, and<br>Farah E.<br>Shamout. [85] | MIMIC-IV; MIM-<br>IC-CXR/(70%train-<br>ing, 10%validation,<br>20%testing)/X-ray,<br>clinical time series data | Flexible fusion framework<br>designed to function without<br>dependence on modality-specific<br>encoders                                                           | LSTM                | Sampling, discretization,<br>standardization                                                                    | ① phenotype classifi-<br>cation<br>② in-hospital mortal-<br>ity prediction | AUC:AUPRC<br>① 0.768; 0.429<br>② 0.874; 0.567                                                                                                                | X                   |
| 2022/Zhou,<br>Tongxue, et al.<br>[86]                                             | BraTS 2018/285<br>(80%training, 20%test-<br>ing)/MRI (T1, FLAIR,<br>T1c and T2)                               | Combine multiple information<br>sources and focus on the miss-<br>ing features via generator                                                                       | U-net               | Cropping, resizing                                                                                              | Segmentation<br>① WT<br>② TC<br>③ ET                                       | DSC<br>① missing T2<br>87.1<br>② missing T2<br>86.6<br>③ missing T2<br>78.2                                                                                  | X                   |
| 2021/Zhu,<br>Yian, et al. [87]                                                    | BraTS<br>2018/(80%training,<br>20%validation)/MRI<br>(T1, FLAIR, T1c and<br>T2)                               | A model that uses two residual<br>paths and multiple variational<br>autoencoders to reconstruct<br>features by adding small differ-<br>ences to the original input | U-net               | Extraction of 2D slices,<br>random normalization,<br>cropping, augmentation<br>(random axis mirror<br>flipping) | Segmentation<br>① WT<br>② TC<br>③ ET                                       | DSC<br>① missing T1 (complete data)<br>87.98 (87.90)<br>② missing T1 (complete data)<br>78.20 (78.82)<br>③ missing T2 (complete data)<br>79.10 (79.19)       | X                   |
| 2021/Zhu,<br>Yian, et al. [88]                                                    | BraTS<br>2018/(80%training,<br>20%validation)/MRI<br>(T1, FLAIR, T1c and<br>T2)                               | Apply one module to generate<br>shared features via cascade<br>operation and another module<br>to fuse the real features with<br>generated ones.                   | U-net               | Extraction of 2D slices,<br>random normalization,<br>cropping, augmentation<br>(random axis mirror<br>flipping) | Segmentation<br>① WT<br>② TC<br>③ ET                                       | DSC<br>① missing T1c (complete<br>data)<br>87.6 (88.3)<br>② missing T2 (complete data)<br>78.2 (77.7)<br>③ missing T2 (complete data)<br>69.3 (68.5)         | X                   |

Table 6  
(Continued)

| Latent feature space-based methods                                           |                                                                                                                                                                                    |                                                                                                                                    |                                        |                                                                                                                                                                                                          |                                             |                                                                                                                                                                                          |
|------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Year/author/<br>citation                                                     | Dataset/no. of<br>subjects/modalities                                                                                                                                              | Main context                                                                                                                       | Architecture                           | Preprocessing                                                                                                                                                                                            | Task                                        | External<br>validation                                                                                                                                                                   |
| 2021/Zhou,<br>Tongxue, et al.<br>[89]                                        | -BraTS 2018/285 training, 66 validation;<br>-BraTS 2019/335 training, 125 validation/<br>MRI (T1, FLAIR, T1c<br>and T2)                                                            | Use a correlation model and fusion strategy based on attention mechanism                                                           | U-net                                  | Co-registration,<br>interpolation,<br>skull-stripping, manual<br>labeling of ground truth,<br>cropping, resizing,<br>distortion correction,<br>normalization                                             | Segmentation<br>① WT<br>② TC<br>③ ET        | DSC<br>① missing T1 (complete data)<br>87.9 (88.2)<br>② missing T1 (complete data)<br>77.5 (78.6)<br>③ missing T2 (complete data)<br>68.4 (69.4)                                         |
|                                                                              |                                                                                                                                                                                    |                                                                                                                                    |                                        |                                                                                                                                                                                                          |                                             | N/A                                                                                                                                                                                      |
| 2021/Huang,<br>Ruobing, et al.<br>[90]                                       | Private dataset A/1560 (1022 training, 100 validation, 438 testing)/ultrasound (B-mode, shear wave elastography, strain elastography, and Doppler)<br>Private dataset B/163/B-mode | Recovery blocks use existing data to rebuild important features of missing modalities at an abstract representation level          | CNN                                    | Cropping, resizing                                                                                                                                                                                       | Classification<br>① breast module diagnosis | ACC; SEN; SPE; Precision (PRE); F1 score (FIS)<br>① (datasetA)<br>Missing Doppler<br>90.65; 91.57; 89.67; 90.66; 91.10<br>(datasetB)<br>Only B-mode<br>79.26; 62.59; 87.52; 72.28; 66.52 |
|                                                                              |                                                                                                                                                                                    |                                                                                                                                    |                                        |                                                                                                                                                                                                          |                                             | N/A                                                                                                                                                                                      |
| 2020/Hu, Dan,<br>et al. [91]                                                 | UNC/UMN Baby Con-nectome Project/178/structural MRI (sMRI), functional (fMRI)                                                                                                      | Latent variable disentanglement strategy and imputation algorithm                                                                  | CNN with adversarial autoencoder [115] | Co-registration, intensity inhomogeneity correction, skull stripping, cerebellum removal, tissue segmentation, hemispheres separation, topological correction, inner/middle/outer surface reconstruction | ① missing imaging modality imputation       | ① Mean relative absolute error (MRAE); RMSE missing: 0.95; 0.677                                                                                                                         |
|                                                                              |                                                                                                                                                                                    |                                                                                                                                    |                                        |                                                                                                                                                                                                          |                                             | N/A                                                                                                                                                                                      |
| 2019/van Gar-<br>deren, Karin,<br>Marion Smits,<br>and Stefan<br>Klein. [92] | BraTS 2018/278 (80%/training, 20%/testing)/MRI (T1, FLAIR, T1c and T2)                                                                                                             | Curriculum learning approach used for fusing information of different modalities                                                   | U-net                                  | Skull-stripping,<br>co-registration,<br>resampling, normaliza-<br>tion                                                                                                                                   | Segmentation<br>① WT<br>② TC<br>③ ET        | DSC<br>① missing T1 (complete data)<br>81 (83)<br>② missing T1 (complete data)<br>64 (71)<br>③ missing T2 (complete data)<br>63 (63)                                                     |
|                                                                              |                                                                                                                                                                                    |                                                                                                                                    |                                        |                                                                                                                                                                                                          |                                             | N/A                                                                                                                                                                                      |
| 2019/Lau,<br>Kenneth, Jonas<br>Adler, and Jens<br>Sjölund. [93]              | BraTS 2018/285 (70%/training, 30%/validation)/MRI (T1, FLAIR, T1c and T2)                                                                                                          | Model trained to handle missing data using information from the available modalities, by training with randomly dropped modalities | U-net                                  | Bias field correction,<br>normalization, split                                                                                                                                                           | Segmentation<br>① WT<br>② TC<br>③ ET        | DSC<br>① missing T1 (complete data)<br>86.0 (86.1)<br>② missing T1 (complete data)<br>78.8 (78.0)<br>③ missing T2 (complete data)<br>72.3 (71.3)                                         |
|                                                                              |                                                                                                                                                                                    |                                                                                                                                    |                                        |                                                                                                                                                                                                          |                                             | N/A                                                                                                                                                                                      |

Table 6  
(Continued)

| Latent feature space-based methods           |                                                                                                                                                                                                                                                                 |                                                                                                                                                                                 |                                                 |                                                                                                                         |                                                  |                                                                                                                                                                                                                                                                                                 |                                                                                                                                               |                        |
|----------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------|------------------------|
| Year/author/<br>citation                     | Dataset/no. of<br>subjects/modalities                                                                                                                                                                                                                           | Main context                                                                                                                                                                    | Architecture                                    | Preprocessing                                                                                                           | Task                                             | Metrics/performance                                                                                                                                                                                                                                                                             | Code                                                                                                                                          | External<br>validation |
| 2019/Dorent,<br>Reuben, et al.<br>[94]       | BraTS 2018/285<br>(70%/training, 10%/val-<br>idation, 20% testing)/<br>MRI (T1, FLAIR, DWI<br>and T2)                                                                                                                                                           | Extend Multi-modal Varia-<br>tional Auto-Encoders for 3D<br>segmentation as well as mixture<br>sampling strategy                                                                | 3D U-net                                        | Standardization<br>(augmentation) random<br>flipping, rotation                                                          | Segmentation<br>① WT<br>② TC<br>③ ET             | DSC<br>① missing T1 (complete data)<br>88.6 (88.8)<br>② missing T1 (complete data)<br>75.6 (76.4)<br>③ missing T1 (complete data)<br>71.2 (71.7)                                                                                                                                                | N/A                                                                                                                                           | X                      |
|                                              |                                                                                                                                                                                                                                                                 |                                                                                                                                                                                 |                                                 |                                                                                                                         |                                                  |                                                                                                                                                                                                                                                                                                 |                                                                                                                                               |                        |
| 2017/Chartsias,<br>Agisilaos, et al.<br>[95] | BraTS 2015/54 (42<br>training, 6 validation,<br>6 testing)/MRI (T1,<br>FLAIR, DWI and T2)<br>ISLES 2015/28 (22<br>training, 3 validation,<br>3 testing)/MRI (T1,<br>FLAIR, T1c and T2)<br>IXI/28 (22 training, 3<br>validation, 3 testing)/<br>MRI (T1, T2, PD) | Model maps all input modalities<br>to a common representation and<br>uses one decoder per output<br>type, guided by a custom loss<br>to ensure consistency across<br>modalities | Fully convolu-<br>tional network<br>(FCN) [116] | Trimming, normalization                                                                                                 | ① synthesized image<br>quality measurement       | Mean squared error (MSE);<br>Structural similarity index<br>measure (SSIM); Peak signal-<br>to-noise ratio (PSNR)<br>① test on ISLES<br>T1toT2<br>0.299; 0.831; 25.78<br>T1toFLAIR<br>0.268; 0.831; 29.99<br>test on BraTS<br>T1toT2<br>0.333; 0.929; 30.96<br>T1toFLAIR<br>0.283; 0.897; 30.32 | <a href="https://github.com/agis85/multimodal_brain_synthesis">https://github.<br/>com/agis85/<br/>multimod-<br/>al_brain_syn-<br/>thesis</a> | X                      |
| 2016/Havaei,<br>Mohammad, et<br>al. [96]     | MSGC/43 (20 training,<br>23 testing)/MRI (T1,<br>FLAIR and T2)<br>RRMS/300/MRI (T1,<br>FLAIR, T1c and T2)<br>BraTS2015/274/MRI<br>(T1, FLAIR, T1c and<br>T2)<br>BraTS2013/35/MRI<br>(T1, FLAIR, T1c and<br>T2)                                                  | Learn a latent space embedding<br>for each modality, enabling<br>arithmetic operations within that<br>space                                                                     | CNN                                             | Bias field correction,<br>intensity normalization,<br>truncation, unit variance,<br>co-registration, interpo-<br>lation | Segmentation<br>① WT<br>② TC<br>③ ET<br>④ lesion | DSC<br>① (BraTS)<br>missing T1c (complete data)<br>83.87 (83.15)<br>② (BraTS)<br>missing T1 (complete data)<br>70.62 (72.5)<br>③ (BraTS)<br>missing T2 (complete data)<br>71.30 (75.37)<br>④ (RRMS)<br>missing T1c (complete data)<br>46.6 (48.66)                                              | N/A                                                                                                                                           | ✓                      |

**Table 7**  
**Evaluation metrics used for evaluation of methods dealing with missing modalities**

| Evaluation metrics                     |                                                                                                                                                                                                                                                                                                                                                                                    |                                                                                                                                                                                                                                                                            |
|----------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Metric                                 | Formulation                                                                                                                                                                                                                                                                                                                                                                        | Usage                                                                                                                                                                                                                                                                      |
| Accuracy (ACC)                         | Given true positive (TP), true negative (TN), positive (P) and negative (N),<br>$ACC = \frac{TP+TN}{P+N}$                                                                                                                                                                                                                                                                          | ACC ranges from 0 to 1 (or in percentage terms), with the larger value showing that the classification result is more precise.                                                                                                                                             |
| Sensitivity (SEN)/ Recall ((REC)       | Given true positive (TP) and positive (P),<br>$SEN = \frac{TP}{P}$                                                                                                                                                                                                                                                                                                                 | SEN ranges from 0 to 1, with the larger value showing better performance in correctly identifying positive cases                                                                                                                                                           |
| Specificity (SPE)                      | Given true negative (TN) and negative (N),<br>$SPE = \frac{TN}{N}$                                                                                                                                                                                                                                                                                                                 | SPE ranges from 0 to 1, with the larger value showing better performance in correctly identifying negative cases                                                                                                                                                           |
| Area under the curve (AUC)             | $AUC = \int_{x=0}^1 TPR(FPR^{-1}(x)) dx$<br>where FPR(T): $T \rightarrow x$ , the x-axis of ROC curve, denotes false positive rate, TPR(T): $T \rightarrow y(x)$ , the y-axis of ROC curve, denotes true positive rate.                                                                                                                                                            | AUC ranges from 0.5 to 1, with the larger value showing that the model is more likely to arrange positive instance in front of negative instance (i.e. has a better ability of prediction).                                                                                |
| Precision (PRE)                        | Given true positive (TP) and false positive (FP),<br>$Precision = \frac{TP}{TP+FP}$                                                                                                                                                                                                                                                                                                | Higher precision values indicate better model performance in correctly identifying positive instances.                                                                                                                                                                     |
| Average precision (AP)                 | $AP = \frac{\sum_{k=1}^n (P(k) \cdot \Delta R(k))}{\sum_{k=1}^n \delta(k)}$<br>where $n$ is the total number of retrieved items, $P(k)$ is the precision at cut-off $k$ , $\Delta R(k)$ represents the change in recall from the previous cut-off to the current cut-off, $\delta(k)$ is an indicator function that is 1 if the item at position $k$ is relevant, and 0 otherwise. | Higher AP values indicate better model performance, with 1 representing perfect precision and recall, and 0 representing no relevant items retrieved.                                                                                                                      |
| F1 score (F1S)                         | Given Precision (PRE) and Recall (REC),<br>$F1 = \frac{2 \times PRE \times REC}{PRE + REC}$                                                                                                                                                                                                                                                                                        | F1 Score is widely used in binary classification tasks, especially when the classes are imbalanced. It ranges from 0 to 1, where 0 indicates poor performance, suggesting either low precision or low recall, 1 indicates perfect precision and recall.                    |
| Dice similarity coefficient (DSC)      | Given two sets A and B<br>$DSC = \frac{2 \times  A \cap B }{ A  +  B }$                                                                                                                                                                                                                                                                                                            | DSC is commonly used image segmentation, where the goal is to compare the similarity between two sets of regions. Higher Dice coefficient values indicate better agreement between the segmented regions and the ground truth, indicating better segmentation performance. |
| Balanced accuracy (BAC)                | Given true positive (TP), false positive (FP), true negative (TN) and false negative (FN),<br>$BAC = \frac{1}{2} \left( \frac{TP}{TP+FN} + \frac{TN}{TN+FP} \right)$                                                                                                                                                                                                               | BAC is used to evaluate the performance of binary classification models, particularly with imbalanced datasets. Higher BAC values indicate better model performance, with 1 representing perfect classification accuracy and 0.5 representing random classification.       |
| Balanced classification accuracy (BCA) | Given Sensitivity (SEN) and Specificity (SPE)<br>$BCA = \frac{1}{2} (SEN + SPE)$                                                                                                                                                                                                                                                                                                   | BCA is also used to evaluate the performance of binary classification models, particularly with imbalanced datasets. Higher BCA values indicate better model performance, with 1 representing perfect classification accuracy and 0 representing random classification.    |
| Matthews correlation coefficient (MCC) | Given true positive (TP), false positive (FP), true negative (TN) and false negative (FN),<br>$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$                                                                                                                                                                                                      | Higher MCC values indicate better classifier performance, with 1 representing perfect agreement between prediction and observation, 0 indicating random prediction, and -1 indicating total disagreement.                                                                  |
| Mean absolute error (MAE)              | $MAE = \frac{1}{n} \sum_{i=1}^n  y_i - \hat{y}_i $<br>where $n$ is the number of samples, $\hat{y}_i$ is sample $i$ 's predicted value, and $y_i$ is sample $i$ 's true value                                                                                                                                                                                                      | MAE is widely used as a measure of the accuracy of regression models. Lower MAE values indicate better model performance, as they suggest smaller errors between predictions and actual values.                                                                            |
| Mean relative absolute error (MRAE)    | $MRAE = \frac{1}{n} \sum_{i=1}^n \left  \frac{y_i - \hat{y}_i}{y_i} \right $<br>where $n$ is the number of samples, $\hat{y}_i$ is sample $i$ 's predicted value, and $y_i$ is sample $i$ 's true value                                                                                                                                                                            | In [91], MRAE is the mean of the absolute error divided by the corresponding chronological age. A lower MRAE indicates better model performance.                                                                                                                           |

Table 7  
(Continued)

| Evaluation metrics                                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                                                                                                                                                                                                                                 |
|---------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Metric                                                  | Formulation                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | Usage                                                                                                                                                                                                                           |
| Mean squared error (MSE)                                | $MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$ <p>where <math>n</math> is the number of samples, <math>\hat{y}_i</math> is sample <math>i</math>'s predicted value, and <math>y_i</math> is sample <math>i</math>'s true value</p>                                                                                                                                                                                                                                                                                                                           | MSE is commonly used for model evaluation, model comparison, and model selection in regression tasks. Lower MSE values indicate better model performance, as they suggest smaller errors between predictions and actual values. |
| Normalized mean squared error (NMSE)                    | $NMSE = \frac{1}{n} \sum_{i=1}^n \frac{(y_i - \hat{y}_i)^2}{(y_i - \bar{y})^2}$ <p>where <math>n</math> is the number of samples, <math>\hat{y}_i</math> is sample <math>i</math>'s predicted value, <math>y_i</math> is sample <math>i</math>'s true value and <math>\bar{y}</math> is the mean of true values.</p>                                                                                                                                                                                                                                               | In [51], NMSE was used to measure the model's ability of image synthesis. A lower NMSE value indicates a better synthesis quality.                                                                                              |
| Root mean squared error (RMSE)                          | $RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$ <p>where <math>n</math> is the number of samples, <math>\hat{y}_i</math> is sample <math>i</math>'s predicted value, and <math>y_i</math> is sample <math>i</math>'s true value</p>                                                                                                                                                                                                                                                                                                                   | In [78], RMSE was used to measure model's ability of imputation. A smaller RMSE value indicates a better imputation performance.                                                                                                |
| Peak signal-to-noise ratio (PSNR)                       | <p>Given the maximum possible pixel value (MAX) and Mean squared error (MSE) between the ground truth and synthesized images,</p> $PSNR = 10 \cdot \log_{10} \left( \frac{MAX^2}{MSE} \right)$                                                                                                                                                                                                                                                                                                                                                                     | PSNR usually ranges from 20 to 50 (dB), with a larger value showing that the synthesized (reconstructed) image's quality is better.                                                                                             |
| Structural similarity index measure (SSIM)              | $SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$ <p>where <math>C_1</math> and <math>C_2</math> are constants to stabilize the division with weak denominator, <math>\mu_x</math> is the mean of image <math>x</math>, <math>\sigma_x^2</math> is the standard deviations of image <math>x</math>, <math>\sigma_{xy}</math> is the covariance of <math>x</math> and <math>y</math></p>                                                                                                      | SSIM value is a decimal between -1 and 1, where 1 indicates perfect similarity, 0 indicates no similarity, and -1 indicates perfect dissimilarity. A higher SSIM value suggests a higher similarity between the images.         |
| Mean structural similarity index measure (MSSIM)        | <p>Given Structural similarity index measure (SSIM),</p> $MSSIM(x, y) = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n SSIM(x_{ij}, y_{ij})$ <p>where <math>x</math> and <math>y</math> are the compared images, <math>m</math> and <math>n</math> are the dimensions of the images, and <math>x_{ij}</math>, <math>y_{ij}</math> are the corresponding patches in images <math>x</math> and <math>y</math>.</p>                                                                                                                                                           | In [52], a higher MSSIM value indicates higher similarity between images.                                                                                                                                                       |
| Learned perceptual image patch similarity (LPIPS) [117] | $LPIPS(x, y) = \frac{1}{N} \sum_{i=1}^N \left( \frac{1}{C} \sum_{c=1}^C \left( \frac{f_c(x_i) - f_c(y_i)}{R} \right)^2 \right)^{1/2}$ <p>where <math>x</math> and <math>y</math> are two images, <math>N</math> is the number of patches, <math>C</math> is the number of channels, <math>f_c</math> is the feature extraction function, <math>R</math> is scaling factor.</p>                                                                                                                                                                                     | In [45], LPIPS was utilized to capture the perceptual similarity between images. A lower LPIPS value indicates improved perceptual quality.                                                                                     |
| Feature similarity index measure (FSIM) [118]           | $FSIM(x, y) = \frac{\sum_{i,j} S(x_i, y_i) \cdot f(x_i, y_i)}{\sum_{i,j} S(x_i, y_i)}$ <p>where <math>x</math> and <math>y</math> are two images, <math>s</math> and <math>f</math> represents the local similarity and the structural feature similarity between corresponding image patches <math>x_i, y_i</math> respectively.</p>                                                                                                                                                                                                                              | In [43], a higher FSIM value indicates higher similarity between the compared images.                                                                                                                                           |
| Maximum mean discrepancy (MMD)                          | $MMD^2(\mathcal{F}, \mathcal{G}) = \left  \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j=1}^n \phi(x_i) - \frac{2}{mn} \sum_{i=1}^n \sum_{j=1}^m \phi(x_i) + \frac{1}{m(m-1)} \sum_{i=1}^m \sum_{j=1}^m \phi(y_i) \right ^2$ <p>where <math>\phi</math> represents for feature map, <math>x_i</math> and <math>y_i</math> are from distribution <math>\mathcal{F}</math> and <math>\mathcal{G}</math> respectively, while <math>n</math> and <math>m</math> are numbers of subjects of distribution <math>\mathcal{F}</math> and <math>\mathcal{G}</math> respectively.</p> | In [53], a lower MMD value indicates better image quality.                                                                                                                                                                      |

Table 7  
(Continued)

| Evaluation metrics        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                                                                                                                        |
|---------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Metric                    | Formulation                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | Usage                                                                                                                                                                                  |
| Concordance index (CI)    | $CI = \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j=1}^N I(\widehat{y}_i > \widehat{y}_j) \cdot \delta(y_i < y_j)$ <p>where <math>N</math> is the total number of samples, <math>\widehat{y}_i</math> is sample <math>i</math>'s predicted value, <math>y_i</math> is sample <math>i</math>'s true value. A pair of <math>i</math> and <math>j</math> is concordant if the predicted ranking agrees with the true ranking (i.e., if <math>\widehat{y}_i &gt; \widehat{y}_j</math> and <math>y_i &lt; y_j</math> and discordant otherwise.</p> | In [83], CI was used to measure how well the model can reliably rank patients based on their predicted survival times. A larger CI, ranging from 0 to 1, indicates better performance. |
| Hausdorff distance (HD95) | <p>Let <math>A</math> and <math>B</math> be two sets of points in a metric space, <math>HD95(A, B) = \max(\text{percentile}(d(a, B), 95), \text{percentile}(d(b, A), 95))</math></p> <p>where <math>d(x, Y)</math> represents the distance between point <math>x</math> from set <math>X</math> and the nearest point in set <math>Y</math>, the function <math>\text{percentile}(x, p)</math> returns the <math>p</math>th percentile of the distances and the 95th percentile is chosen to calculate the maximum distance.</p>      | In [74], HD 95 was used to evaluate each nested subregion of brain tumors. A lower HD95 indicate better segmentation performance.                                                      |

Table 8

Quantitative summary of the three methodological categories. Columns indicate (I) typical tasks, (II) median performance (DSC for segmentation, AUC for classification, SSIM for synthesis), (III) robustness to missing modalities, (IV) use of external validation, and (V) code availability

| Category                   | Typical tasks                                  | Median reported performance                                                         | Robustness                                            | External validation | Code availability |
|----------------------------|------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------|---------------------|-------------------|
| Image synthesis methods    | Mostly synthesis; segmentation; classification | DSC: ~0.86 (Segmentation)<br>AUC: ~0.85 (Classification)<br>SSIM: ~0.93 (Synthesis) | Limited (mainly tested for specific missing modality) | More frequently     | Low (~10%)        |
| Knowledge transfer methods | Mostly segmentation; classification            | DSC: ~0.88 (Segmentation)<br>AUC: ~0.84 (Classification)                            | Moderate (can handle missing modalities)              | Rarely              | High (~70%)       |
| Latent space-based methods | Mostly segmentation; classification            | DSC: ~0.86 (Segmentation)<br>AUC: ~0.83 (Classification)                            | Moderate (can handle missing modalities)              | Rarely              | Moderate (~30%)   |

segmentation tasks [65, 71–76], and 42% are performed on classification tasks [66–70]. All the methods for segmentation purpose choose U-net, while for classification the choice of network ranges from CNNs to GAN. The difficulty of designing the network, the high training cost of the source model as well as the interpretability of the model have always been the major concern concerns for knowledge transfer methods [119]. Even though none of the reviewed papers completely overcame these hindrances, some of them made notable progress. In [71] authors disentangled features on the modality level and employed a contrastive learning-based learning scheme in the spatial and frequency domain to exploit more explicit relations between modalities. Similarly in [66], a contrastive learning-based loss was applied to direct the optimization of the target model from the source model. The authors in [68] combined image synthesis models with domain adaptation by implementing two GAN to generate missing modalities and transfer knowledge within the learned feature space.

Thirdly, our results show that 33% of the reviewed papers use latent feature space-based methods. Among them, 55% are implemented for the purpose of segmentation [77, 81, 84, 86–89, 92–94, 96], 20% are implemented for classification tasks [79, 80, 85, 90], and the rest are

implemented for various uses, such as prediction [78, 82, 83], missing modality imputation [91] and synthesized image evaluation [95]. Of the reviewed latent feature space-based methods, 80% rely on convolution neural networks (CNNs) to extract latent features. Specifically, around 50% of them use U-net, which has a comprehensible structure and performs well on the segmentation tasks [120]. Notably 10% of the reviewed papers leveraged transformers to achieve satisfactory outcomes [79, 84]. Considering that the bias of convolution limits the ability to harness cross-modal relationships [86], a multi-modal transformer can model the correlated high-level features from different modalities by the attention mechanism [110]. Moreover, the transformer has inspired researchers to combine the attention mechanism with the traditional encoder–decoder model, enabling the network to highlight important latent features while suppressing irrelevant ones [77]. The remaining 10% of the methods use RNN-based blocks in their network when dealing with longitudinal missing modality [78, 82]. Multi-modal data fusion strategies are utilized by most latent feature space-based methods to combine the multi-modal extracted features into latent space representation. They contribute to exploiting the latent feature space and emphasizing the major features from multiple modalities.

The fusion strategies based on arithmetic operations such as merging the features by computing mean and variance are simple but effective when the purpose is to reinforce the model's robustness. More complex fusion strategies, including L1 or L2 distance minimization [95] and attention-based fusion [89], have also been proposed. However, it is difficult to decide which strategy is best for handling multi-modal missing imaging modality since none of them assure the model to learn a shared latent representation from different intensity distributions of different modalities. Interpretability of the multi-modal deep learning models is another concern for their further applications in medical image analysis [121]. Consequently, researchers are striving to enhance the interpretability of their models. For example, in [81] authors not only presented a latent feature space-based U-net effective for both classification and segmentation tasks but also enhanced the interpretability via the t-SNE visualization of the latent feature space.

Although a formal meta-analysis was not feasible due to the substantial heterogeneity across studies in terms of datasets, tasks, and performance metrics, in Table 8, we provide a quantitative summary of the three methodological categories, comparing their typical applications, reported performance, robustness to missing modalities, external validation, and code availability. Across the three categories, reported performance is generally similar. Segmentation tasks, which are mostly evaluated on the BraTS dataset, show DSC values around 0.86–0.88, while classification tasks on the ADNI dataset reach AUC values of about 0.83–0.85. For image synthesis, SSIM is typically around 0.93. The main differences lie in robustness and research practice: image synthesis methods are usually tested on specific missing modalities, show more frequent external validation, but have the lowest code availability. In contrast, knowledge transfer and latent space methods are more often tested in missing modalities scenarios, though they are less often externally validated, with code availability being high (~70%) and moderate (~30%), respectively.

While the above categories provide a useful framework to organize the literature, they are not fully separate. Some studies use methods that overlap across categories. For example, certain works generate synthetic images to expand datasets while also applying knowledge transfer to improve model performance [68]. Others combine latent feature space analysis with image synthesis to strengthen prediction tasks [36]. These mixed approaches show that methods in this field are often built on each other, and that combining techniques from different categories can be especially useful for tackling the challenges of multimodal medical imaging.

In the review, we have summarized the datasets used for missing modality studies. Regarding the most popular datasets used to develop missing modalities techniques, 60% of the approaches included in this review used the BraTS dataset and 29% selected the ADNI dataset for training. Most datasets contain multiple MRI modalities, resulting in similar preprocessing pipelines. It is worth noting that public datasets partly reflect real-world clinical practice. They are derived from real patients and usually include preprocessing steps such as normalization or artifact reduction, which are also common in practice. This makes them useful for benchmarking and method development. However, they are not fully representative, as the released data are often high-quality and relatively clean. Cases with noise, missing data, or poor annotations are usually filtered out. Therefore, results based only on public datasets may not fully capture the challenges of everyday clinical scenarios.

With respect to experimental design, preprocessing was the most frequently discussed aspect. For MRI preprocessing, normalization, registration, and skull stripping are the top three frequently used methods, while for other modalities such as PET, registration and normalization remained standard practice. In terms of computational resources, 90% of the papers listed in this review used Graphical Processing Units

(GPUs) for their training. Only 25% of the reviewed papers provided a public code, half of which are latent feature space-based methods, promoting reproducibility and collaboration for advancing studies. There are 42% reviewed papers using multiple datasets, 23% of which perform an external validation [38, 40, 50, 52, 55, 59], which helps to identify biases or variability present in the training samples while strengthening the robustness and dependability of their models.

Lastly, a detailed summary of all evaluation metrics used by the papers included in this review is provided in Section 3.3. In most studies, the chosen assessment criteria depend on the task categories. Around 34% of papers included in this review evaluate their methods under one missing modality condition. Among the rest of papers tested on multiple missing modalities scenario, half of them carry out 'robustness evaluation' [122], i.e. assessing their networks on every possible combination of missing modalities. For non-synthesis methods, it is difficult to quantify how well the recovery is because in some categories, it is not even possible to verify the quality of the recovered modality. For image synthesis methods, such evaluation is available due to the presence of ground truth. However, it is hard to compare the recovered modalities within different image synthesis models even if they are applied to the same task and the same dataset, if they synthesize different imaging modalities. Also, since most image synthesis-based methods use CNN networks that rely on fixed channels of input and output, it is difficult for them to perform a robust evaluation as a separate model will be required for each possible input–output scenario [45]. In contrast, nearly half of the articles on latent feature space-based methods and knowledge transfer methods implement the 'robustness evaluation'.

## 5. Limitation

Firstly, the review was conducted by only two reviewers. Although cross-checking was carried out, the small number of reviewers may have introduced some subjectivity in study selection and data extraction. Secondly, the included studies are more weighted toward computer science and methodological aspects, while clinical perspectives may be underrepresented. This could limit the assessment of clinical relevance. Thirdly, it was difficult to compare results across studies in a quantitative way, since they used different datasets, preprocessing steps, evaluation metrics, and tasks. The lack of standardized benchmarks makes direct comparison challenging. Nevertheless, given the systematic search across multiple major databases and the structured synthesis process, we are confident that this review covered the most relevant studies and that the findings capture the current landscape of research on deep learning methods for missing modalities in multimodal medical imaging.

## 6. Future Direction

Despite the rapid development of deep learning methods, there is still a long way to go to solve the missing modalities problem perfectly. In this section, we will discuss the limitations of current approaches and indicate the possible solutions concerning them.

A major challenge for adopting deep learning-based models in a clinical-setting is posed by that lack of model interpretability. For clinical decision-making applications, model interpretability and introspection are crucial components. It is known that features extracted from deep learning models are abstract and may not always be clinically relevant [6]. Such models will not generalize well on unseen data or be biased against certain populations. However, the interpretation of deep learning models is an active research topic and multiple methods have been proposed [123–127] that allow us to visualize which parts of the data the model considers important for its predictions, even though the

underlying feature representations remain abstract. Ref. [123] shows a popular deep learning technique that visually identifies the key areas within an input image influencing the model's predictions. It is beneficial in shedding light on how decisions are made by multi-modal deep learning models applied to medical images. Emerging solutions aim not only to enhance interpretability but also to uncover associations between different modalities [6, 128].

From the perspective of multi-modal data dimensionality, we observe in our results that 2D is a more popular choice than 3D. Although 3D medical image analysis offers advantages in terms of capturing richer spatial information and potentially improving prediction accuracy, it comes with the increased complexity of models. For instance, CNNs used for 3D image analysis may require deeper architecture or incorporate additional layers to handle the data. Adapting transfer learning is a potential solution to this challenge. Pre-trained models from related domains will help the 3D target model to initialize more efficiently and reduce the need for extensive labeled data [129]. Another noteworthy approach is to integrate attention mechanisms into 3D models. By focusing on the important parts and ignoring irrelevant regions of the 3D images, the attention modules can reduce the computation cost and improve the interpretability of the 3D models [130].

Another important future direction for successfully dealing with missing modalities will be gaining data diversity. The BraTS and the ADNI datasets provide a large sum of image data for the reviewed papers' experiment; however, most of the data are of MRI modalities, bringing inadequate data diversity issues. As it is difficult to collect various and complete imaging data from clinical scenarios, generated images might be a better way to enrich the data diversity. To achieve such a purpose, more research on the fidelity of the synthesized images and the interpretability of image synthesis methods are required. Other modalities of data such as clinical records or genomic information are also applicable to expand the data variety and improve the effectiveness of the models.

In addition to data diversity, another critical consideration is the lack of consistent ways to evaluate methods across studies. While various papers report quantitative metrics, differences in tasks, datasets, and evaluation metrics make it hard to compare results directly. Therefore, future work should create standardized benchmarks and shared evaluation protocols. This would make cross-study comparisons more meaningful, improve reproducibility, and provide reliable baselines to guide future research.

Despite the popularity of deep learning methods, a major roadblock for its widespread adoption is the unavailability of large-scale groundtruth datasets with all multimodal data. This challenge arises primarily from the labor-intensive process of manual annotation, privacy concerns in the clinical domain, and biases introduced by the typically small size of patient cohorts, which often represent high socioeconomic status [131, 132]. Popular approaches like data augmentation [133], semi-supervised learning [134], transfer learning [135], and automated annotation [136] offer promising solutions to address the problem of insufficient labeled data.

## 7. Conclusion

This review identified and summarized 61 relevant papers through the search process in accordance with the PRISMA guidelines. Our research is distinct in its focus on systematically reviewing multi-modal deep learning methods developed over the last decade dealing with missing imaging modality issues in medical image analysis. Central to our investigation are key research inquiries answered in the results section, illustrating the methodologies employed in medical image

analysis with missing modality, cataloging publicly available datasets for researchers, and encapsulating prevalent evaluation techniques. We provide a detailed discussion regarding our findings and identify notable research gaps and their potential solutions. The literature explores addressed questions and analyzes them in the following sections, highlighting a rapidly growing and globally significant field of interest.

## Funding Support

This work is supported in part by the National Institutes of Health (NIH, R01 EB029944, R01-EB030582) and the Academic and Research Committee Awards of Cincinnati Children's Hospital Medical Center.

## Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

## Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

## Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

## Author Contribution Statement

**Jinzhao Qian:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Ankita Joshi:** Conceptualization, Methodology, Validation, Investigation, Writing – original draft, Writing – review & editing, Visualization, Supervision. **Hailong Li:** Writing – review & editing, Project administration. **Nehal A. Parikh:** Writing – review & editing. **Jonathan R. Dillman:** Writing – review & editing. **Lili He:** Conceptualization, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

## References

- [1] Muhammad, G., Alshehri, F., Karray, F., El Saddik, A., Alsulaiman, M., & Falk, T. H. (2021). A comprehensive survey on multimodal medical signals fusion for smart healthcare systems. *Information Fusion*, 76, 355–375. <https://doi.org/10.1016/j.inffus.2021.06.007>
- [2] Azam, M. A., Khan, K. B., Salahuddin, S., Rehman, E., Khan, S. A., Khan, M. A., ... & Gandomi, A. H. (2022). A review on multimodal medical image fusion: Compendious analysis of medical modalities, multimodal databases, fusion techniques and quality metrics. *Computers in Biology and Medicine*, 144, 105253. <https://doi.org/10.1016/j.combiomed.2022.105253>
- [3] Venugopalan, J., Tong, L., Hassanzadeh, H. R., & Wang, M. D. (2021). Multimodal deep learning models for early detection of Alzheimer's disease stage. *Scientific Reports*, 11(1), 3254. <https://doi.org/10.1038/s41598-020-74399-w>
- [4] Ourselin, S., Stefanescu, R., & Pennec, X. (2002). Robust registration of multi-modal images: Towards real-time clinical applications. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 140–147. [https://doi.org/10.1007/3-540-45787-9\\_18](https://doi.org/10.1007/3-540-45787-9_18)

- [5] Liu, S., Cai, W., Liu, S., Zhang, F., Fulham, M., Feng, D., ... & Kikinis, R. (2015). Multimodal neuroimaging computing: A review of the applications in neuropsychiatric disorders. *Brain Informatics*, 2(3), 167–180. <https://doi.org/10.1007/s40708-015-0019-x>
- [6] Lipkova, J., Chen, R. J., Chen, B., Lu, M. Y., Barbieri, M., Shao, D., ... & Mahmood, F. (2022). Artificial intelligence for multimodal data integration in oncology. *Cancer Cell*, 40(10), 1095–1110. <https://doi.org/10.1016/j.ccell.2022.09.012>
- [7] Stahlschmidt, S. R., Ulfenborg, B., & Synnergren, J. (2022). Multimodal deep learning for biomedical data fusion: A review. *Briefings in Bioinformatics*, 23(2), bbab569. <https://doi.org/10.1093/bib/bbab569>
- [8] Tay, W. L., Chui, C. K., Ong, S. H., & Ng, A. C. M. (2013). Ensemble-based regression analysis of multimodal medical data for osteopenia diagnosis. *Expert Systems with Applications*, 40(2), 811–819. <https://doi.org/10.1016/j.eswa.2012.08.031>
- [9] Van der Heijden, G. J., Donders, A. R. T., Stijnen, T., & Moons, K. G. (2006). Imputation of missing values is superior to complete case analysis and the missing-indicator method in multivariable diagnostic research: A clinical example. *Journal of Clinical Epidemiology*, 59(10), 1102–1109. <https://doi.org/10.1016/j.jclinepi.2006.01.015>
- [10] de Goeij, M. C., Van Diepen, M., Jager, K. J., Tripepi, G., Zoccali, C., & Dekker, F. W. (2013). Multiple imputation: Dealing with missing data. *Nephrology Dialysis Transplantation*, 28(10), 2415–2420. <https://doi.org/10.1093/ndt/gft221>
- [11] Brand, L., Nichols, K., Wang, H., Shen, L., & Huang, H. (2019). Joint multi-modal longitudinal regression and classification for Alzheimer's disease prediction. *IEEE Transactions on Medical Imaging*, 39(6), 1845–1855. <https://doi.org/10.1109/TMI.2019.2958943>
- [12] Ma, M., Ren, J., Zhao, L., Tulyakov, S., Wu, C., & Peng, X. (2021). Smil: Multimodal learning with severely missing modality. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(3), 2302–2310. <https://doi.org/10.1609/aaai.v35i3.16330>
- [13] Li, M., Jiang, Y., Zhang, Y., & Zhu, H. (2023). Medical image analysis using deep learning algorithms. *Frontiers in Public Health*, 11, 1273253. <https://doi.org/10.3389/fpubh.2023.1273253>
- [14] Li, J., Hong, D., Gao, L., Yao, J., Zheng, K., Zhang, B., & Chanussot, J. (2022). Deep learning in multimodal remote sensing data fusion: A comprehensive review. *International Journal of Applied Earth Observation and Geoinformation*, 112, 102926. <https://doi.org/10.1016/j.jag.2022.102926>
- [15] Xie, F., Yuan, H., Ning, Y., Ong, M. E. H., Feng, M., Hsu, W., ... & Liu, N. (2022). Deep learning for temporal data representation in electronic health records: A systematic review of challenges and methodologies. *Journal of Biomedical Informatics*, 126, 103980. <https://doi.org/10.1016/j.jbi.2021.103980>
- [16] Ehle, A., Rindtorff, N. T., Brinker, T. J., Luedde, T., Pearson, A. T., & Kather, J. N. (2021). Deep learning in cancer pathology: A new generation of clinical biomarkers. *British Journal of Cancer*, 124(4), 686–696. <https://doi.org/10.1038/s41416-020-01122-x>
- [17] Azad, R., Dehghanmanshadi, M., Khosravi, N., Cohen-Adad, J., & Merhof, D. (2025). Addressing missing modality challenges in MRI images: A comprehensive review. *Computational Visual Media*, 11(2), 241–268. <https://doi.org/10.26599/CVM.2025.9450399>
- [18] Biratu, E. S., Schwenker, F., Ayano, Y. M., & Debelee, T. G. (2021). A survey of brain tumor segmentation and classification algorithms. *Journal of Imaging*, 7(9), 179. <https://doi.org/10.3390/jimaging7090179>
- [19] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372. <https://doi.org/10.1136/bmj.n71>
- [20] Jack Jr, C. R., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., ... & Weiner, M. W. (2008). The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *Journal of Magnetic Resonance Imaging*, 27(4), 685–691. <https://doi.org/10.1002/jmri.21049>
- [21] Baid, U., Ghodasara, S., Mohan, S., Bilello, M., Calabrese, E., Colak, E., ... & Bakas, S. (2023). *RSNA-ASNR-MICCAI-BraTS-2021* [Data set]. National Cancer Institute. <https://doi.org/10.7937/JC8X-9874>
- [22] Maier, O., Menze, B. H., Von der Gabelntz, J., Häni, L., Heinrich, M. P., Liebrand, M., ... & Reyes, M. (2017). ISLES 2015-A public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI. *Medical Image Analysis*, 35, 250–269. <https://doi.org/10.1016/j.media.2016.07.009>
- [23] Marcus, D. S., Wang, T. H., Parker, J., Csernansky, J. G., Morris, J. C., & Buckner, R. L. (2007). Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. *Journal of Cognitive Neuroscience*, 19(9), 1498–1507. <https://doi.org/10.1162/jocn.2007.19.9.1498>
- [24] Johnson, A. E., Pollard, T. J., Berkowitz, S. J., Greenbaum, N. R., Lungren, M. P., Deng, C. Y., ... & Horng, S. (2019). MIMIC-CXR, a de-identified publicly available database of chest radiographs with free-text reports. *Scientific Data*, 6(1), 317. <https://doi.org/10.1038/s41597-019-0322-0>
- [25] Marek, K., Jennings, D., Lasch, S., Siderowf, A., Tanner, C., Simuni, T., ... & Parkinson Progression Marker Initiative. (2011). The Parkinson progression marker initiative (PPMI). *Progress in Neurobiology*, 95(4), 629–635. <https://doi.org/10.1016/j.pneurobio.2011.09.005>
- [26] Shattuck, D. W., Mirza, M., Adisetiyo, V., Hojatkashani, C., Salamon, G., Narr, K. L., ... & Toga, A. W. (2008). Construction of a 3D probabilistic atlas of human cortical structures. *Neuroimage*, 39(3), 1064–1080. <https://doi.org/10.1016/j.neuroimage.2007.09.031>
- [27] Valindria, V. V., Pawlowski, N., Rajchl, M., Lavdas, I., Aboagye, E. O., Rockall, A. G., ... & Glocker, B. (2018). Multi-modal learning from unpaired images: Application to multi-organ segmentation in CT and MRI. In *2018 IEEE Winter Conference on Applications of Computer Vision*, 547–556. <https://doi.org/10.1109/WACV.2018.00066>
- [28] Kavur, A. E., Gezer, N. S., Barış, M., Aslan, S., Conze, P. H., Groza, V., ... & Selver, M. A. (2021). CHAOS challenge-combined (CT-MR) healthy abdominal organ segmentation. *Medical Image Analysis*, 69, 101950. <https://doi.org/10.1016/j.media.2020.101950>
- [29] Gu, D., Shi, F., Hua, R., Wei, Y., Li, Y., Zhu, J., ... & Consortium of Chinese Brain Molecular and Functional Mapping (CBMFM). (2023). An artificial-intelligence-based age-specific template construction framework for brain structural analysis using magnetic resonance images. *Human Brain Mapping*, 44(3), 861–875. <https://doi.org/10.1002/hbm.26126>
- [30] Gillman, M. W., & Blaisdell, C. J. (2018). Environmental influences on child health outcomes, a research program of the National Institutes of Health. *Current Opinion in Pediatrics*, 30(2), 260–262. <https://doi.org/10.1097/MOP.0000000000000600>
- [31] Styner, M., Lee, J., Chin, B., Chin, M., Commowick, O., Tran, H., ... & Warfield, S. (2008). 3D segmentation in the clinic: A grand

- challenge II: MS lesion segmentation. *MIDAS Journal*, 2008, 1–6.
- [32] Xiao, S., Lewis, M., Mellor, D., McCabe, M., Byrne, L., Wang, T., ... & Dong, S. (2016). The China longitudinal ageing study: Overview of the demographic, psychosocial and cognitive data of the Shanghai sample. *Journal of Mental Health*, 25(2), 131–136. <https://doi.org/10.3109/09638237.2015.1124385>
- [33] Ellis, K. A., Bush, A. I., Darby, D., De Fazio, D., Foster, J., Hudson, P., ... & AIBL Research Group. (2009). The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's disease. *International Psychogeriatrics*, 21(4), 672–687. <https://doi.org/10.1017/S1041610209009405>
- [34] Howell, B. R., Styner, M. A., Gao, W., Yap, P. T., Wang, L., Baluyot, K., ... & Ellison, J. T. (2019). The UNC/UMN Baby Connectome Project (BCP): An overview of the study design and protocol development. *NeuroImage*, 185, 891–905. <https://doi.org/10.1016/j.neuroimage.2018.03.049>
- [35] Bullitt, E., Zeng, D., Gerig, G., Aylward, S., Joshi, S., Smith, J. K., ... & Ewend, M. G. (2005). Vessel tortuosity and brain tumor malignancy: A blinded study1. *Academic Radiology*, 12(10), 1232–1240. <https://doi.org/10.1016/j.acra.2005.05.027>
- [36] Kebaili, A., Lapuyade-Lahorgue, J., Vera, P., & Ruan, S. (2025). AMM-Diff: Adaptive multi-modality diffusion network for missing modality imputation. In *2025 IEEE 22nd International Symposium on Biomedical Imaging*, 1–4. <https://doi.org/10.1109/ISBI60581.2025.10980985>
- [37] Guo, J., Bu, R., Shen, W., & Feng, T. (2025). Towards robust multimodal ultrasound classification for liver tumor diagnosis: A generative approach to modality missingness. *Computer Methods and Programs in Biomedicine*, 265, 108759. <https://doi.org/10.1016/j.cmpb.2025.108759>
- [38] Deng, Y., Ling, J., Rao, X., Tan, J., Fu, X., & Li, S. (2024). An improved Pix2Pix GAN for medical image generation. In *IAPR Workshop on Artificial Neural Networks in Pattern Recognition*, 99–110. [https://doi.org/10.1007/978-3-031-71602-7\\_9](https://doi.org/10.1007/978-3-031-71602-7_9)
- [39] Zhang, M., Sun, L., Kong, Z., Zhu, W., Yi, Y., & Yan, F. (2024). Pyramid-attentive GAN for multimodal brain image complementation in Alzheimer's disease classification. *Biomedical Signal Processing and Control*, 89, 105652. <https://doi.org/10.1016/j.bspc.2023.105652>
- [40] Gao, X., Shi, F., Shen, D., & Liu, M. (2023). Multimodal transformer network for incomplete image generation and diagnosis of Alzheimer's disease. *Computerized Medical Imaging and Graphics*, 110, 102303. <https://doi.org/10.1016/j.compmedimag.2023.102303>
- [41] Wang, T., Wang, H., Deng, J., Zhang, D., Feng, J., & Chen, B. (2023). Feature generation and multi-sequence fusion based deep convolutional network for breast tumor diagnosis with missing MR sequences. *Biomedical Signal Processing and Control*, 82, 104536. <https://doi.org/10.1016/j.bspc.2022.104536>
- [42] Jiang, Y., Zhang, S., & Chi, J. (2023). Multi-modal brain tumor data completion based on reconstruction consistency loss. *Journal of Digital Imaging*, 36(4), 1794–1807. <https://doi.org/10.1007/s10278-022-00697-6>
- [43] Cao, B., Bi, Z., Hu, Q., Zhang, H., Wang, N., Gao, X., & Shen, D. (2023). Autoencoder-driven multimodal collaborative learning for medical image synthesis. *International Journal of Computer Vision*, 131(8), 1995–2014. <https://doi.org/10.1007/s11263-023-01791-0>
- [44] Wang, Y., Hu, H., Yu, S., Yang, Y., Guo, Y., Song, X., ... & Liu, Q. (2023). A unified hybrid transformer for joint MRI sequences super-resolution and missing data imputation. *Physics in Medicine & Biology*, 68(13), 135006. <https://doi.org/10.1088/1361-6560/acdc80>
- [45] Liu, J., Pasumarthi, S., Duffy, B., Gong, E., Datta, K., & Zaharchuk, G. (2023). One model to synthesize them all: Multi-contrast multi-scale transformer for missing data imputation. *IEEE Transactions on Medical Imaging*, 42(9), 2577–2591. <https://doi.org/10.1109/TMI.2023.3261707>
- [46] Wu, J., Guo, D., Wang, L., Yang, S., Zheng, Y., Shapey, J., ... & Wang, G. (2023). TISS-net: Brain tumor image synthesis and segmentation using cascaded dual-task networks and error-prediction consistency. *Neurocomputing*, 544, 126295. <https://doi.org/10.1016/j.neucom.2023.126295>
- [47] Jin, L., Zhao, K., Zhao, Y., Che, T., & Li, S. (2022). A hybrid deep learning method for early and late mild cognitive impairment diagnosis with incomplete multimodal data. *Frontiers in Neuroinformatics*, 16, 843566. <https://doi.org/10.3389/fninf.2022.843566>
- [48] Sun, Y., Liu, Y., & Liu, B. (2022). Predicting conversion to mild cognitive impairment in cognitively normal with incomplete multi-modal neuroimages. In *2022 10th International Conference on Bioinformatics and Computational Biology*, 61–65. <https://doi.org/10.1109/ICBCB55259.2022.9802479>
- [49] Zhang, J., He, X., Qing, L., Gao, F., & Wang, B. (2022). BP-GAN: Brain PET synthesis from MRI using generative adversarial network for multi-modal Alzheimer's disease diagnosis. *Computer Methods and Programs in Biomedicine*, 217, 106676. <https://doi.org/10.1016/j.cmpb.2022.106676>
- [50] Liu, Y., Yue, L., Xiao, S., Yang, W., Shen, D., & Liu, M. (2022). Assessing clinical progression from subjective cognitive decline to mild cognitive impairment with incomplete multi-modal neuroimages. *Medical Image Analysis*, 75, 102266. <https://doi.org/10.1016/j.media.2021.102266>
- [51] Huang, P., Li, D., Jiao, Z., Wei, D., Cao, B., Mo, Z., ... & Shen, D. (2022). Common feature learning for brain tumor MRI synthesis by context-aware generative adversarial network. *Medical Image Analysis*, 79, 102472. <https://doi.org/10.1016/j.media.2022.102472>
- [52] Kaplan, S., Perrone, A., Alexopoulos, D., Kenley, J. K., Barch, D. M., Buss, C., ... & Smyser, C. D. (2022). Synthesizing pseudo-T2w images to recapture missing data in neonatal neuroimaging with applications in rs-fMRI. *Neuroimage*, 253, 119091. <https://doi.org/10.1016/j.neuroimage.2022.119091>
- [53] Gao, X., Shi, F., Shen, D., & Liu, M. (2021). Task-induced pyramid and attention GAN for multimodal brain image imputation and classification in Alzheimer's disease. *IEEE Journal of Biomedical and Health Informatics*, 26(1), 36–43. <https://doi.org/10.1109/JBHI.2021.3097721>
- [54] Yan, K., Liu, Z., Zheng, S., Guo, Z., Zhu, Z., & Zhao, Y. (2021). Coarse-to-fine learning framework for semi-supervised multimodal MRI synthesis. In *Asian conference on pattern recognition*, 370–384. [https://doi.org/10.1007/978-3-031-02444-3\\_28](https://doi.org/10.1007/978-3-031-02444-3_28)
- [55] Islam, M., Wijethilake, N., & Ren, H. (2021). Glioblastoma multiforme prognosis: MRI missing modality generation, segmentation and radiogenomic survival prediction. *Computerized Medical Imaging and Graphics*, 91, 101906. <https://doi.org/10.1016/j.compmedimag.2021.101906>
- [56] Akbar, M. U., Murino, V., & Sona, D. (2021). Multimodal segmentation of medical images with heavily missing data. In *2021 IEEE EMBS International Conference on Biomedical and Health Informatics*, 1–4. <https://doi.org/10.1109/BHI50953.2021.9508609>
- [57] Lin, W., Lin, W., Chen, G., Zhang, H., Gao, Q., Huang, Y., ... & Alzheimer's Disease Neuroimaging Initiative. (2021). Bidirec-

- tional mapping of brain MRI and PET with 3D reversible GAN for the diagnosis of Alzheimer's disease. *Frontiers in Neuroscience*, 15, 646013. <https://doi.org/10.3389/fnins.2021.646013>
- [58] Pan, Y., Liu, M., Lian, C., Xia, Y., & Shen, D. (2020). Spatially-constrained fisher representation for brain disease identification with incomplete multi-modal neuroimages. *IEEE Transactions on Medical Imaging*, 39(9), 2965–2975. <https://doi.org/10.1109/TMI.2020.2983085>
- [59] Liu, Y., Pan, Y., Yang, W., Ning, Z., Yue, L., Liu, M., & Shen, D. (2020). Joint neuroimage synthesis and representation learning for conversion prediction of subjective cognitive decline. In *International Conference on Medical Image Computing and Computer Assisted Intervention*, 583–592. [https://doi.org/10.1007/978-3-030-59728-3\\_57](https://doi.org/10.1007/978-3-030-59728-3_57)
- [60] Dar, S. U., Yurt, M., Shahdloo, M., Ildiz, M. E., Tınaz, B., & Çukur, T. (2020). Prior-guided image reconstruction for accelerated multi-contrast MRI via generative adversarial networks. *IEEE Journal of Selected Topics in Signal Processing*, 14(6), 1072–1087. <https://doi.org/10.1109/JSTSP.2020.3001737>
- [61] Cao, B., Zhang, H., Wang, N., Gao, X., & Shen, D. (2020). Auto-GAN: Self-supervised collaborative learning for medical image synthesis. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(7), 10486–10493. <https://doi.org/10.1609/aaai.v34i07.6619>
- [62] Sharma, A., & Hamarneh, G. (2019). Missing MRI pulse sequence synthesis using multi-modal generative adversarial network. *IEEE Transactions on Medical Imaging*, 39(4), 1170–1183. <https://doi.org/10.1109/TMI.2019.2945521>
- [63] Pan, Y., Liu, M., Lian, C., Zhou, T., Xia, Y., & Shen, D. (2018). Synthesizing missing PET from MRI with cycle-consistent generative adversarial networks for Alzheimer's disease diagnosis. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 455–463. [https://doi.org/10.1007/978-3-030-00931-1\\_52](https://doi.org/10.1007/978-3-030-00931-1_52)
- [64] Tang, Z., Yap, P. T., & Shen, D. (2018). A new multi-atlas registration framework for multimodal pathological images using conventional monomodal normal atlases. *IEEE Transactions on Image Processing*, 28(5), 2293–2304. <https://doi.org/10.1109/TIP.2018.2884563>
- [65] Liu, W., Hou, J., Zhong, X., Zhan, H., Cheng, J., Fang, Y., & Yue, G. (2025). Improving multi-modal brain tumor segmentation via pre-training and knowledge distillation based post-training. *Neurocomputing*, 640, 130318. <https://doi.org/10.1016/j.neucom.2025.130318>
- [66] Xing, X., Zhu, M., Chen, Z., & Yuan, Y. (2024). Comprehensive learning and adaptive teaching: Distilling multi-modal knowledge for pathological glioma grading. *Medical Image Analysis*, 91, 102990. <https://doi.org/10.1016/j.media.2023.102990>
- [67] Chen, Y., Pan, Y., Xia, Y., & Yuan, Y. (2023). Disentangle first, then distill: A unified framework for missing modality imputation and Alzheimer's disease diagnosis. *IEEE Transactions on Medical Imaging*, 42(12), 3566–3578. <https://doi.org/10.1109/TMI.2023.3295489>
- [68] Dolci, G., Rahaman, M. A., Galazzo, I. B., Cruciani, F., Abrol, A., Chen, J., ... & Calhoun, V. D. (2023). Deep generative transfer learning predicts conversion to Alzheimer's disease from neuroimaging genomics data. In *IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops*, 1–5. <https://doi.org/10.1109/ICASSPW59220.2023.10193683>
- [69] Dolci, G., Rahaman, M. A., Chen, J., Duan, K., Fu, Z., Abrol, A., ... & Calhoun, V. D. (2022). A deep generative multi-modal imaging genomics framework for Alzheimer's disease prediction. In *2022 IEEE 22nd International Conference on Bioinformatics and Bioengineering*, 41–44. <https://doi.org/10.1109/BIBE55377.2022.00017>
- [70] Jeong, S. W., Cho, H. H., Lee, S., & Park, H. (2022). Robust multimodal fusion network using adversarial learning for brain tumor grading. *Computer Methods and Programs in Biomedicine*, 226, 107165. <https://doi.org/10.1016/j.cmpb.2022.107165>
- [71] Yang, Q., Guo, X., Chen, Z., Woo, P. Y., & Yuan, Y. (2022). D 2-Net: Dual disentanglement network for brain tumor segmentation with missing modalities. *IEEE Transactions on Medical Imaging*, 41(10), 2953–2964. <https://doi.org/10.1109/TMI.2022.3175478>
- [72] Li, H., Li, C., Huang, W., Zheng, X., Xi, Y., & Wang, S. (2023). DIGEST: Deeply supervised knowledge transfer network learning for brain tumor segmentation with incomplete multi-modal MRI scans. In *2023 IEEE 20th International Symposium on Biomedical Imaging*, 1–4. <https://doi.org/10.1109/ISBI53787.2023.10230344>
- [73] Vadachchino, S., Mehta, R., Sepahvand, N. M., Nichyporuk, B., Clark, J. J., & Arbel, T. (2021). Had-net: A hierarchical adversarial knowledge distillation network for improved enhanced tumour segmentation without post-contrast images. In *Medical Imaging with Deep Learning*, 143, 787–801.
- [74] Wang, Y., Zhang, Y., Liu, Y., Lin, Z., Tian, J., Zhong, C., ... & He, Z. (2021). ACN: Adversarial co-training network for brain tumor segmentation with missing modalities. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 410–420. [https://doi.org/10.1007/978-3-030-87234-2\\_39](https://doi.org/10.1007/978-3-030-87234-2_39)
- [75] Hu, M., Maillard, M., Zhang, Y., Ciceri, T., La Barbera, G., Bloch, I., & Gori, P. (2020). Knowledge distillation from multi-modal to mono-modal segmentation networks. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 772–781. [https://doi.org/10.1007/978-3-030-59710-8\\_75](https://doi.org/10.1007/978-3-030-59710-8_75)
- [76] Shen, Y., & Gao, M. (2019). Brain tumor segmentation on MRI with missing modalities. In *International Conference on Information Processing in Medical Imaging*, 417–428. [https://doi.org/10.1007/978-3-030-20351-1\\_32](https://doi.org/10.1007/978-3-030-20351-1_32)
- [77] Zhou, T. (2023). Feature fusion and latent feature learning guided brain tumor segmentation and missing modality recovery network. *Pattern Recognition*, 141, 109665. <https://doi.org/10.1016/j.patcog.2023.109665>
- [78] Wang, T., Chen, X., Zhang, X., Zhou, S., Feng, Q., & Huang, M. (2023). Multi-view imputation and cross-attention network based on incomplete longitudinal and multimodal data for conversion prediction of mild cognitive impairment. *Expert Systems with Applications*, 231, 120761. <https://doi.org/10.1016/j.eswa.2023.120761>
- [79] Gao, X., Shi, F., Shen, D., & Liu, M. (2023). Feature-Based transformer with incomplete multimodal brain images for diagnosis of neurodegenerative diseases. In *International Workshop on Predictive Intelligence in Medicine*, 25–34. [https://doi.org/10.1007/978-3-031-46005-0\\_3](https://doi.org/10.1007/978-3-031-46005-0_3)
- [80] Wang, T., Chen, X., Zhang, J., Feng, Q., & Huang, M. (2023). Deep multimodality-disentangled association analysis network for imaging genetics in neurodegenerative diseases. *Medical Image Analysis*, 88, 102842. <https://doi.org/10.1016/j.media.2023.102842>
- [81] Wang, H., Chen, Y., Ma, C., Avery, J., Hull, L., & Carneiro, G. (2023). Multi-modal learning with missing modality via shared-specific feature modelling. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 15878–15887.

- [82] Morar, U., Martin, H., M, R. P., Izquierdo, W., Zarafshan, E., Forouzaneshad, P., ... & Adjouadi, M. (2023). Prediction of cognitive test scores from variable length multimodal data in Alzheimer's disease. *Cognitive Computation*, 15(6), 2062–2086. <https://doi.org/10.1007/s12559-023-10169-w>
- [83] Hou, W., Lin, C., Yu, L., Qin, J., Yu, R., & Wang, L. (2023). Hybrid graph convolutional network with online masked autoencoder for robust multimodal cancer survival prediction. *IEEE Transactions on Medical Imaging*, 42(8), 2462–2473. <https://doi.org/10.1109/TMI.2023.3253760>
- [84] Zhang, Y., He, N., Yang, J., Li, Y., Wei, D., Huang, Y., ... & Zheng, Y. (2022). mmformer: Multimodal medical transformer for incomplete multimodal learning of brain tumor segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 107–117. [https://doi.org/10.1007/978-3-031-16443-9\\_11](https://doi.org/10.1007/978-3-031-16443-9_11)
- [85] Hayat, N., Geras, K. J., & Shamout, F. E. (2022). MedFuse: Multi-modal fusion with clinical time-series data and chest X-ray images. In *7th Machine Learning for Healthcare Conference*, 479–503.
- [86] Zhou, T., Vera, P., Canu, S., & Ruan, S. (2022). Missing data imputation via conditional generator and correlation learning for multimodal brain tumor segmentation. *Pattern Recognition Letters*, 158, 125–132. <https://doi.org/10.1016/j.patrec.2022.04.019>
- [87] Zhu, Y., Wang, S., Hu, Y., Ma, X., Qin, Y., & Xie, J. (2021). DRM-VAE: A dual residual multi variational auto-encoder for brain tumor segmentation with missing modalities. In *2021 IEEE 4th International Conference on Electronics and Communication Engineering*, 82–86. <https://doi.org/10.1109/ICECE54449.2021.9674673>
- [88] Zhu, Y., Wang, S., Lin, R., Hu, Y., & Chen, Q. (2021). Brain tumor segmentation for missing modalities by supplementing missing features. In *2021 IEEE 6th International Conference on Cloud Computing and Big Data Analytics*, 652–656. <https://doi.org/10.1109/ICCCBDA51879.2021.9442533>
- [89] Zhou, T., Canu, S., Vera, P., & Ruan, S. (2021). Latent correlation representation learning for brain tumor segmentation with missing MRI modalities. *IEEE Transactions on Image Processing*, 30, 4263–4274. <https://doi.org/10.1109/TIP.2021.3070752>
- [90] Huang, R., Lin, Z., Dou, H., Wang, J., Miao, J., Zhou, G., ... & Ni, D. (2021). AW3M: An auto-weighting and recovery framework for breast cancer diagnosis using multi-modal ultrasound. *Medical Image Analysis*, 72, 102137. <https://doi.org/10.1016/j.media.2021.102137>
- [91] Hu, D., Zhang, H., Wu, Z., Wang, F., Wang, L., Smith, J. K., ... & Shen, D. (2020). Disentangled-multimodal adversarial autoencoder: Application to infant age prediction with incomplete multimodal neuroimages. *IEEE Transactions on Medical Imaging*, 39(12), 4137–4149. <https://doi.org/10.1109/TMI.2020.3013825>
- [92] van Garderen, K., Smits, M., & Klein, S. (2019). Multi-modal segmentation with missing MR sequences using pre-trained fusion networks. In *MICCAI Workshop on Domain Adaptation and Representation Transfer*, 165–172. [https://doi.org/10.1007/978-3-030-33391-1\\_19](https://doi.org/10.1007/978-3-030-33391-1_19)
- [93] Lau, K., Adler, J., & Sjölund, J. (2019). A unified representation network for segmentation with missing modalities. *arXiv Preprint:1908.06683*. <https://doi.org/10.48550/arXiv.1908.06683>
- [94] Dorent, R., Joutard, S., Modat, M., Ourselin, S., & Vercauteren, T. (2019). Hetero-modal variational encoder-decoder for joint modality completion and segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 74–82. [https://doi.org/10.1007/978-3-030-32245-8\\_9](https://doi.org/10.1007/978-3-030-32245-8_9)
- [95] Chartsias, A., Joyce, T., Giuffrida, M. V., & Tsaftaris, S. A. (2017). Multimodal MR synthesis via modality-invariant latent representation. *IEEE Transactions on Medical Imaging*, 37(3), 803–814. <https://doi.org/10.1109/TMI.2017.2764326>
- [96] Havai, M., Guizard, N., Chapados, N., & Bengio, Y. (2016). Hemis: Hetero-modal image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 469–477. [https://doi.org/10.1007/978-3-319-46723-8\\_54](https://doi.org/10.1007/978-3-319-46723-8_54)
- [97] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, 63(11), 139–144. <https://doi.org/10.1145/3422622>
- [98] Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE International Conference on Computer Vision*, 2223–2232.
- [99] Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. A. (2018). Generative adversarial networks: An overview. *IEEE Signal Processing Magazine*, 35(1), 53–65. <https://doi.org/10.1109/MSP.2017.2765202>
- [100] Saad, M. M., Rehmani, M. H., & O'Reilly, R. (2022). Addressing the intra-class mode collapse problem using adaptive input image normalization in GAN-based X-ray images. In *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society*, 2049–2052. <https://doi.org/10.1109/EMBC48229.2022.9871260>
- [101] Vincent, P., Larochelle, H., Bengio, Y., & Manzagol, P. A. (2008). Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th International Conference on Machine Learning*, 1096–1103. <https://doi.org/10.1145/1390156.1390294>
- [102] Jung, E., Luna, M., & Park, S. H. (2021). Conditional GAN with an attention-based generator and a 3D discriminator for 3D medical image generation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 318–328. [https://doi.org/10.1007/978-3-030-87231-1\\_31](https://doi.org/10.1007/978-3-030-87231-1_31)
- [103] Qin, D., Bu, J. J., Liu, Z., Shen, X., Zhou, S., Gu, J. J., ... & Dai, H. F. (2021). Efficient medical image segmentation based on knowledge distillation. *IEEE Transactions on Medical Imaging*, 40(12), 3820–3831. <https://doi.org/10.1109/TMI.2021.3098703>
- [104] Guan, H., & Liu, M. (2021). Domain adaptation for medical image analysis: a survey. *IEEE Transactions on Biomedical Engineering*, 69(3), 1173–1185. <https://doi.org/10.1109/TBME.2021.3117407>
- [105] Qin, X., Zhang, Z., Huang, C., Dehghan, M., Zaiane, O. R., & Jagersand, M. (2020). U2-Net: Going deeper with nested U-structure for salient object detection. *Pattern Recognition*, 106, 107404. <https://doi.org/10.1016/j.patcog.2020.107404>
- [106] Donders, A. R. T., Van Der Heijden, G. J., Stijnen, T., & Moons, K. G. (2006). A gentle introduction to imputation of missing values. *Journal of Clinical Epidemiology*, 59(10), 1087–1091. <https://doi.org/10.1016/j.jclinepi.2006.01.014>
- [107] Santos, M. S., Pereira, R. C., Costa, A. F., Soares, J. P., Santos, J., & Abreu, P. H. (2019). Generating synthetic missing data: A review by missing mechanism. *IEEE Access*, 7, 11651–11667. <https://doi.org/10.1109/ACCESS.2019.2891360>

- [108] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 234–241. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
- [109] Nguyen, M., He, T., An, L., Alexander, D. C., Feng, J., Yeo, B. T., & Alzheimer's Disease Neuroimaging Initiative. (2020). Predicting Alzheimer's disease progression using deep recurrent neural networks. *NeuroImage*, 222, 117203. <https://doi.org/10.1016/j.neuroimage.2020.117203>
- [110] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30.
- [111] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778.
- [112] Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). 3D U-Net: Learning dense volumetric segmentation from sparse annotation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 424–432. [https://doi.org/10.1007/978-3-319-46723-8\\_49](https://doi.org/10.1007/978-3-319-46723-8_49)
- [113] Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. *Neural Computation*, 31(7), 1235–1270. [https://doi.org/10.1162/neco\\_a\\_01199](https://doi.org/10.1162/neco_a_01199)
- [114] Yan, M., Deng, L., Hu, X., Liang, L., Feng, Y., Ye, X., ... & Xie, Y. (2020). Hygen: A GCN accelerator with hybrid architecture. In *2020 IEEE International Symposium on High Performance Computer Architecture*, 15–29. <https://doi.org/10.1109/HPCA47549.2020.00012>
- [115] Zhang, H., Guo, W., Zhang, S., Lu, H., & Zhao, X. (2022). Unsupervised deep anomaly detection for medical images using an improved adversarial autoencoder. *Journal of Digital Imaging*, 35(2), 153–161. <https://doi.org/10.1007/s10278-021-00558-8>
- [116] Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3431–3440.
- [117] Zhang, R., Isola, P., Efros, A. A., Shechtman, E., & Wang, O. (2018). The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 586–595.
- [118] Zhang, L., Zhang, L., Mou, X., & Zhang, D. (2011). FSIM: A feature similarity index for image quality assessment. *IEEE Transactions on Image Processing*, 20(8), 2378–2386. <https://doi.org/10.1109/TIP.2011.2109730>
- [119] Xing, X., Hou, Y., Li, H., Yuan, Y., Li, H., & Meng, M. Q. H. (2021). Categorical relation-preserving contrastive knowledge distillation for medical image classification. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 163–173. [https://doi.org/10.1007/978-3-030-87240-3\\_16](https://doi.org/10.1007/978-3-030-87240-3_16)
- [120] Siddique, N., Sidike, P., Elkin, C., & Devabhaktuni, V. (2020). U-Net and its variants for medical image segmentation: Theory and applications. *IEEE Access*, 9, 82031–82057. <https://doi.org/10.1109/ACCESS.2021.3086020>
- [121] Zhu, Q., Xu, B., Huang, J., Wang, H., Xu, R., Shao, W., & Zhang, D. (2023). Deep multi-modal discriminative and interpretability network for Alzheimer's disease diagnosis. *IEEE Transactions on Medical Imaging*, 42(5), 1472–1483. <https://doi.org/10.1109/TMI.2022.3230750>
- [122] Ferreira, S. L., Caires, A. O., Borges, T. D. S., Lima, A. M., Silva, L. O., & dos Santos, W. N. (2017). Robustness evaluation in analytical methods optimized using experimental designs. *Microchemical Journal*, 131, 163–169. <https://doi.org/10.1016/j.microc.2016.12.004>
- [123] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE International Conference on Computer Vision*, 618–626.
- [124] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30.
- [125] Dickinson, Q., & Meyer, J. G. (2022). Positional SHAP (PoSHAP) for Interpretation of machine learning models trained from biological sequences. *PLOS Computational Biology*, 18(1), e1009736. <https://doi.org/10.1371/journal.pcbi.1009736>
- [126] Schaumburg, A. J., Juarez-Nicanor, W. C., Choudhury, S. J., Pastrián, L. G., Pritt, B. S., Prieto Pozuelo, M., ... & Fuchs, T. J. (2020). Interpretable multimodal deep learning for real-time pan-tissue pan-disease pathology search on social media. *Modern Pathology*, 33(11), 2169–2185. <https://doi.org/10.1038/s41379-020-0540-1>
- [127] Begoli, E., Bhattacharya, T., & Kusnezov, D. (2019). The need for uncertainty quantification in machine-assisted medical decision making. *Nature Machine Intelligence*, 1(1), 20–23. <https://doi.org/10.1038/s42256-018-0004-1>
- [128] Chen, R. J., Lu, M. Y., Williamson, D. F., Chen, T. Y., Lipkova, J., Noor, Z., ... & Mahmood, F. (2022). Pan-cancer integrative histology-genomic analysis via multimodal deep learning. *Cancer Cell*, 40(8), 865–878. <https://doi.org/10.1016/j.ccell.2022.07.004>
- [129] Zhao, Z., Alzubaidi, L., Zhang, J., Duan, Y., & Gu, Y. (2024). A comparison review of transfer learning and self-supervised learning: Definitions, applications, advantages and limitations. *Expert Systems with Applications*, 242, 122807. <https://doi.org/10.1016/j.eswa.2023.122807>
- [130] Zhou, H. Y., Guo, J., Zhang, Y., Han, X., Yu, L., Wang, L., & Yu, Y. (2023). nnformer: Volumetric medical image segmentation via a 3D transformer. *IEEE Transactions on Image Processing*, 32, 4036–4045. <https://doi.org/10.1109/TIP.2023.3293771>
- [131] Sharrocks, K., Spicer, J., Camidge, D. R., & Papa, S. (2014). The impact of socioeconomic status on access to cancer clinical trials. *British Journal of Cancer*, 111(9), 1684–1687. <https://doi.org/10.1038/bjc.2014.108>
- [132] Niranjana, S. J., Wenzel, J. A., Martin, M. Y., Fouad, M. N., Vickers, S. M., Konety, B. R., & Durant, R. W. (2021). Perceived institutional barriers among clinical and research professionals: Minority participation in oncology clinical trials. *JCO Oncology Practice*, 17(5), e666–e675. <https://doi.org/10.1200/OP.20.00970>
- [133] Ke, X., Zou, J., & Niu, Y. (2019). End-to-end automatic image annotation based on deep CNN and multi-label data augmentation. *IEEE Transactions on Multimedia*, 21(8), 2093–2106. <https://doi.org/10.1109/TMM.2019.2895511>
- [134] Wu, F., Wang, Z., Zhang, Z., Yang, Y., Luo, J., Zhu, W., & Zhuang, Y. (2015). Weakly semi-supervised deep learning for multi-label image annotation. *IEEE Transactions on Big Data*, 1(3), 109–122. <https://doi.org/10.1109/TBDATA.2015.2497270>
- [135] Yin, S., & Bi, J. (2018). Medical image annotation based on deep transfer learning. In *2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, 47–49. [https://doi.org/10.1109/Cybermatics\\_2018.2018.00042](https://doi.org/10.1109/Cybermatics_2018.2018.00042)

- [136] Krenzer, A., Makowski, K., Hekalo, A., Fitting, D., Troya, J., Zoller, W. G., ... & Puppe, F. (2022). Fast machine learning annotation in the medical domain: A semi-automated video annotation tool for gastroenterologists. *BioMedical Engineering OnLine*, 21(1), 33. <https://doi.org/10.1186/s12938-022-01001-x>

**How to Cite:** Qian, J., Joshi, A., Li, H., Parikh, N. A., Dillman, J. R., & He, L. (2025). Utility of Deep Learning to Address Missing Modalities from Multi-Modal Medical Imaging Studies: A Systematic Review. *Artificial Intelligence and Applications*. <https://doi.org/10.47852/bonviewAIA52026392>