A Survey of Missing Modality Scenarios and Machine Learning in Cardiovascular Healthcare

www.surveyx.cn

Abstract

This survey explores the transformative role of machine learning in addressing the challenges of missing modality scenarios in cardiovascular healthcare. By integrating diverse medical data, machine learning models enhance diagnostic accuracy and improve patient outcomes. However, their integration into clinical workflows is hindered by compatibility issues with existing systems. The survey highlights advancements such as the MMP approach, which improves robustness in missing modality scenarios, and innovative models like DF-DM and DAPA, which demonstrate high accuracy and reliability in multimodal data processing. Despite the promise of deep learning in photoacoustic imaging, significant validation work is required for clinical applicability. Multimodal fusion methods outperform single-modality approaches, enhancing individualized affective experience prediction. Explainable AI systems, such as xAI-EWS, provide high predictive performance and transparency, crucial for clinical adoption. The Flex-MoE model further validates its robustness in handling diverse modality combinations. The study underscores the need for incorporating demographic variables to improve fairness and accuracy in diagnostic algorithms. Emphasizing the importance of standardization, the survey aligns with findings on advancing Surgical Data Science. Future research should focus on enhancing AI explainability, improving user interfaces, and investigating long-term impacts on clinician-patient relationships. Improved data governance frameworks are essential for ensuring reliable, equitable, and effective machine learning applications in cardiovascular healthcare.

1 Introduction

1.1 Significance of Integrating Diverse Medical Data

Integrating diverse medical data types is essential for improving cardiovascular healthcare outcomes through a comprehensive approach to diagnosis and treatment planning. The complexity of cardiovascular diseases necessitates the amalgamation of various data modalities—clinical, laboratory, imaging, and genomic data—to achieve accurate diagnoses and effective interventions. This holistic perspective enhances clinical decision-making [1]. Additionally, the integration of video and audio data, as illustrated in stroke detection, allows for a more thorough patient assessment, resulting in better outcomes [2]. Such integrative strategies are vital for addressing challenges related to data sparsity and the disconnect between online and in-person healthcare IT systems [3].

1.2 Challenges of Missing Modality Scenarios

The integration of diverse medical data modalities in cardiovascular healthcare faces significant challenges, particularly when specific modalities are absent or incomplete. A notable issue is the underperformance of deep learning models trained solely on infrared images, which lack the semantic richness of optical images, especially in difficult environments [4]. Current methodologies often fail to effectively utilize the complementary interactions between different modalities, leading

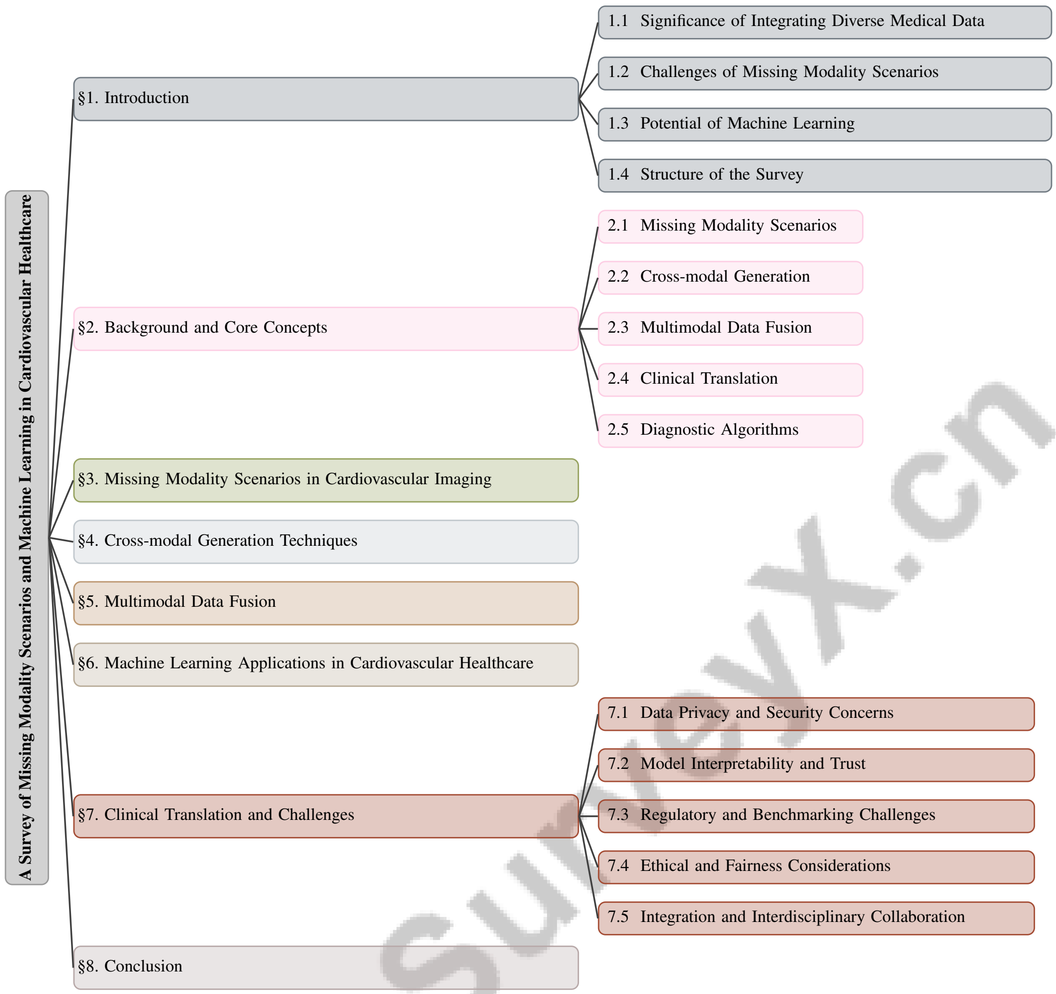


Figure 1: chapter structure

to performance degradation when modalities are missing during inference [5]. This problem is exacerbated by subtle differences in Hounsfield units among blood, thrombus, and arterial walls, which existing methods do not exploit effectively for visualization [6].

In clinical practice, the absence of certain modalities can significantly impact critical processes, such as brain tumor segmentation, where only partial imaging may be available due to degradation and motion artifacts [7]. The high misdiagnosis rate of acute ischemic stroke during emergency triage illustrates the limitations of current methods in detecting subtle stroke signs [2]. Additionally, the complexity of healthcare systems, variability in patient responses, and challenges in accurately modeling these interactions present further obstacles [8]. The reliance on Gaussian assumptions and linear correlations in existing methods fails to accommodate the non-Gaussian distributions and nonlinear dependencies prevalent in many real-world scenarios, thus hindering effective multimodal data fusion [9]. This limitation results in performance degradation in multimodal learning when input modalities are missing during testing [10].

These challenges are compounded by fragmented data systems, lack of integration across platforms, and biases in AI algorithms that can lead to inequitable healthcare outcomes, particularly for marginalized groups [11]. Researchers encounter difficulties in managing the heterogeneous nature of different modalities, including issues such as missing modalities, unpaired data, and the integration of diverse features into a cohesive learning framework [12]. Furthermore, the absence of longitudinal datasets, the complexity of large multimodal imaging examinations, and the need for detailed annotations for traditional supervised machine learning add to the challenges [13].

Practical limitations, such as hand-eye synchronization during ultrasound procedures [1] and the dependency of diagnostic accuracy on the operator's training in stress perfusion CMR [14], further complicate modality integration. Additionally, limited imaging depth in existing optical imaging modalities due to strong scattering in biological tissues presents a significant challenge [15]. The reliance on static models for artery walls and the difficulty in achieving full cycle segmentations restrict the accuracy of hemodynamic assessments [16].

Addressing missing modality scenarios is crucial for effective cardiovascular healthcare. Innovative strategies for managing and integrating incomplete data modalities are essential for enhancing the reliability and accuracy of healthcare outcomes, particularly in multimodal machine learning applications. As healthcare increasingly incorporates artificial intelligence and machine learning to improve clinical decision-making, tackling the challenges posed by missing data becomes imperative. Current research emphasizes the effectiveness of frameworks like M3Care, which utilizes auxiliary patient information to compensate for absent modalities, and the Holistic AI in Medicine (HAIM) framework, demonstrating that leveraging diverse data sources—such as tabular, time-series, textual, and imaging data—can significantly enhance predictive performance across various healthcare tasks. These approaches pave the way for more robust and clinically applicable AI solutions in healthcare [17, 18, 19].

1.3 Potential of Machine Learning

Machine learning (ML) presents substantial promise in addressing the challenges associated with missing modality scenarios in cardiovascular healthcare. By employing advanced algorithms, ML can effectively synthesize missing data modalities from existing information, thereby enhancing diagnostic accuracy and predictive capabilities. Techniques such as feature disentanglement, which separates modality-specific appearance codes from modality-invariant content codes, have demonstrated improvements in segmentation processes even when some modalities are absent [7]. This underscores ML's capacity to manage incomplete datasets while maintaining robust performance in clinical applications.

In stroke detection, the DeepStroke framework exemplifies the application of ML in analyzing multimodal data, specifically facial and speech inputs, to enhance detection accuracy [2]. Such applications highlight ML's ability to integrate diverse data types, compensating for missing modalities and improving clinical decision-making.

Reinforcement learning (RL) within ML frameworks further optimizes treatment protocols and enhances patient outcomes by addressing the complex nature of healthcare decision-making [8]. Additionally, ML-driven methods are being developed to automate quantitative myocardial perfusion evaluations, overcoming technical challenges such as motion correction and labor-intensive processing, thus streamlining diagnostic procedures [14].

ML frameworks also incorporate strategies to ensure fairness and equity in healthcare delivery. For example, fair machine learning practices systematically identify and mitigate biases in healthcare data and models, ensuring equitable outcomes across diverse patient demographics [11]. The OncoNet algorithm, which utilizes weak supervision, exemplifies ML's capability to assess treatment responses effectively, even in the presence of incomplete data [13].

1.4 Structure of the Survey

This survey is meticulously structured to provide a comprehensive exploration of the intersection between missing modality scenarios and machine learning in cardiovascular healthcare. The introductory section establishes the foundational significance of integrating diverse medical data, elucidating the challenges posed by missing modalities and the transformative potential of machine learning to address these issues. Following this, the survey delves into the background and core concepts, offering a detailed examination of missing modality scenarios, cross-modal generation, multimodal data fusion, and their clinical translation within cardiovascular imaging.

The third section focuses on the prevalence and impact of missing data modalities in cardiovascular imaging, scrutinizing existing methods for handling such scenarios and identifying opportunities for improvement. The subsequent section explores cross-modal generation techniques, emphasizing the role of generative models in synthesizing missing modalities. This is followed by an in-depth analysis of multimodal data fusion strategies, including advanced fusion techniques and the incorporation of explainable AI to ensure transparency in diagnostic processes.

Machine learning applications in cardiovascular healthcare are thoroughly reviewed in the sixth section, highlighting deep learning techniques, predictive analytics, and the importance of interpretability and robustness in model development. The penultimate section addresses the challenges of clinical translation, discussing data privacy, model interpretability, regulatory hurdles, ethical considerations, and the necessity for interdisciplinary collaboration.

The survey concludes by synthesizing key findings and insights, emphasizing the vital importance of innovative machine learning applications in addressing the complexities associated with missing modality scenarios. These advancements are crucial for enhancing cardiovascular healthcare outcomes, leveraging multimodal data integration to improve predictive accuracy and clinical decision-making, ultimately leading to more effective patient management and treatment strategies [18, 17, 20].The following sections are organized as shown in Figure 1.

2 Background and Core Concepts

2.1 Missing Modality Scenarios

In cardiovascular imaging, missing modality scenarios occur due to incomplete or absent critical data types, essential for comprehensive clinical evaluations. Factors like sensor malfunctions, patient noncompliance, and privacy issues contribute to these gaps, significantly affecting diagnostic accuracy, especially in intensive care settings [12, 16]. A notable challenge is modeling dynamic artery wall deformations during the cardiac cycle in 4D flow MRI, compounded by missing data [16]. Integrating structurally diverse modalities, such as histopathology images and genomic data, requires preserving each modality's unique characteristics while enabling effective cross-modal interactions, a task complicated by the scarcity of annotated data needed for supervised machine learning [12].

The complexity of detecting signals in high-dimensional multimodal data, while considering intermodal dependencies, further complicates the integration process. Efficiently fusing multimodal datasets, often hindered by incomplete data or resolution variations, can lead to suboptimal patient outcomes. Additionally, correlating heterogeneous data from different modalities due to inconsistent distributions and representations poses a significant hurdle [12]. Addressing these challenges is crucial for enhancing diagnostic accuracy and improving patient outcomes in cardiovascular healthcare. Developing robust embedding spaces from multimodal data, despite missing modalities, is essential for simulating real-world scenarios and facilitating effective cross-modal interactions. Innovative strategies for managing and integrating incomplete data modalities are vital for achieving reliable healthcare outcomes and ensuring equity across diverse patient populations [12].

2.2 Cross-modal Generation

Cross-modal generation techniques address data gaps by leveraging available modalities to infer or reconstruct absent ones, enhancing the robustness of clinical evaluations. These methods are crucial in medical imaging, where datasets are often incomplete due to sensor malfunctions or privacy concerns [21]. The Missing Modality Imagination Network (MMIN) exemplifies this by predicting representations of missing modalities, improving emotion recognition robustness [22]. Frameworks like MultiVI employ probabilistic approaches that incorporate uncertainty in multimodal data integration, enhancing prediction accuracy and reliability [23]. The AutoPrognosis-M framework integrates structured clinical data with medical imaging, reflecting the holistic approach clinicians use in decision-making, highlighting the importance of cross-modal generation in clinical practice [24].

Generative models, particularly those based on convolutional neural networks (CNNs), streamline classification processes in medical imaging, such as classifying pixels in OCT images into distinct tissue types [25]. These models facilitate the reconstruction of missing modalities by exploiting relationships among available data, thus enhancing the overall quality of medical diagnoses. Despite their potential, challenges in integrating modalities due to noise and data fusion complexities can lead to inaccuracies [26]. Categorizing existing research into cross-modal generation techniques provides insights into generating representations in one modality based on another, enriching the understanding of intermodal relationships [27]. This understanding is crucial for developing robust embedding spaces that can simulate real-world scenarios and facilitate effective cross-modal interactions, even with missing modalities [12].

Cross-modal generation techniques are essential for addressing incomplete data challenges in medical imaging, particularly in enhancing the automatic generation of radiology reports. These methods align visual and textual information, extracting relevant insights from medical images despite missing modalities. Approaches such as reinforcement learning for cross-modal memory alignment and contrastive attention models have significantly improved report accuracy, even amidst data biases and complexities in interpreting subtle abnormalities [28, 29, 30, 31, 32]. Consequently, these techniques alleviate radiologists' workload and contribute to more reliable clinical decision-making by ensuring critical information is accurately captured in medical reports.

2.3 Multimodal Data Fusion

Multimodal data fusion is vital in medical diagnostics, integrating diverse data sources to enhance diagnostic accuracy and reliability [33]. By synthesizing information from multiple modalities—such as imaging, genomic, and physiological data—multimodal fusion provides a comprehensive perspective on patient health, addressing single-modality limitations. This integrative strategy improves diagnostic outcomes by leveraging complementary information across data types.

The significance of multimodal data fusion is evident in applications like hyperspectral image classification, where it overcomes performance bottlenecks by delivering rich complementary information from supplementary modalities [34]. In stroke detection, the DeepStroke framework illustrates the potential of multimodal fusion by integrating video and audio data to enhance diagnostic processes [2]. Advanced fusion strategies, such as those employed in CSK-Net, utilize spectral knowledge distillation and mixed feature exchange mechanisms to enhance semantic segmentation performance, demonstrating the adaptability of fusion techniques in both multimodal and missing modality scenarios [4].

Despite these advancements, challenges persist in integrating multimodal data, particularly concerning data heterogeneity and alignment issues [33]. The survey by Bayoudh et al. addresses various aspects of multimodal data fusion, including representation, alignment, and transfer learning, underscoring ongoing efforts to overcome these challenges [26].

2.4 Clinical Translation

Translating research findings into clinical practice, especially in cardiovascular healthcare, involves a multifaceted process that integrates innovative imaging modalities and machine learning models into routine workflows. This process requires a structured framework to ensure that technological advancements yield tangible patient benefits. A critical aspect of this translation is evaluating and comparing various imaging modalities based on their effectiveness, safety, and application areas, as each modality has unique strengths and weaknesses that must be considered in clinical settings [35].

Integrating machine learning models into clinical practice necessitates a deep understanding of the healthcare ecosystem as a complex adaptive system, emphasizing the need for interactions within the system [36]. This understanding is vital for developing models that align with clinical workflows and address practical needs. For example, the development of Single-Input Polarization-Sensitive Optical Coherence Tomography (SIPS) simplifies hardware requirements, facilitating broader clinical use and overcoming key barriers to the application of advanced imaging techniques [37].

Standardization of evaluation frameworks is crucial in clinical translation, providing benchmarks for performance assessment and facilitating the integration of intelligent systems in healthcare [38]. These frameworks ensure that machine learning models yield reliable and interpretable outcomes that clinicians can trust, thereby enhancing their applicability in clinical practice. User-independent assessments play a significant role in ensuring consistent and objective evaluations [14].

The interpretability of machine learning models is another critical factor for clinical adoption, particularly in applications such as predicting cardiac resynchronization therapy (CRT) response from cardiac magnetic resonance images. The lack of transparency in these models poses a significant barrier to clinician trust and model deployment [39]. Addressing this challenge involves developing models that provide clear insights into their decision-making processes, facilitating clinician acceptance and integration into routine practice.

Ethical considerations and data integrity are paramount in applying machine learning in clinical practice. Adhering to the FAIR principles (Findable, Accessible, Interoperable, Reusable) is essential for promoting transparency and accountability in data sharing and maintenance [40]. This approach is crucial for ensuring equitable healthcare outcomes across diverse patient demographics.

Interdisciplinary collaboration is vital for successfully translating machine learning models into clinical practice. Ongoing communication among researchers, clinicians, and policymakers ensures that models are tailored to clinical workflows and meet the practical needs of healthcare providers [38]. This collaboration is essential for addressing regulatory considerations, which play a crucial role in clinical translation, ensuring the safety and efficacy of emerging technologies.

2.5 Diagnostic Algorithms

Diagnostic algorithms are crucial for processing and interpreting cardiovascular data, leveraging advanced machine learning techniques to enhance diagnostic accuracy and clinical decision-making. These algorithms integrate diverse data sources, including patient demographics, clinical variables, and imaging data, to provide comprehensive insights into cardiovascular health. Such integration allows for dynamic predictions of patient needs and efficient allocation of healthcare resources, thereby enhancing overall care quality [41].

The development of diagnostic algorithms often employs logistic regression and other machine learning models to analyze complex datasets, facilitating predictions of conditions such as diabetes in cardiovascular patients [42]. Rigorous evaluation of these models is essential to ensure their relevance and effectiveness in clinical settings. Objective, task-based evaluations are crucial for assessing AI methods in medical imaging, ensuring alignment with clinical outcomes

In recent years, the field of cardiovascular imaging has faced significant challenges related to the handling of missing modalities. These challenges not only hinder diagnostic accuracy but also limit the potential for comprehensive patient assessments. As illustrated in Figure 2, the figure elucidates both the challenges and opportunities associated with addressing these missing modalities. It highlights the limitations of existing methods while simultaneously showcasing potential improvements through cost reduction, advancements in data collection, and the integration of innovative artificial intelligence solutions. This dual focus on challenges and opportunities provides a nuanced understanding of the current landscape and paves the way for future research directions in the domain.

3 Missing Modality Scenarios in Cardiovascular Imaging

3.1 Existing Methods for Handling Missing Modalities

Current strategies for managing missing modalities in cardiovascular imaging face challenges due to their limited ability to integrate and model the diverse data distributions inherent in multimodal datasets. Traditional methods often involve complex processing that can introduce errors, especially in characterizing features such as superficial plaques in intracoronary optical coherence tomography (OCT) imaging [25]. These methods also typically require custom adaptation for each modality combination, which becomes impractical as the number of modalities increases [10].

In polarization-sensitive imaging, the dual-input PS-OCT systems are hindered by complex and costly hardware requirements, limiting their routine clinical use [37]. Furthermore, traditional saliency methods like Grad-CAM and Eigen-CAM have shown limitations with large datasets, particularly in extensive coronary angiogram studies [43], highlighting the need for more advanced techniques. The integration of optical and infrared modalities is challenging as existing methods often fail to maintain performance when one modality is absent, reducing diagnostic accuracy [4]. This is exacerbated by the inherent sparsity and complexity of healthcare data, which complicates the development of effective multimodal solutions [8].

Efforts to enhance fairness in feature selection aim to improve model performance across diverse demographic groups and mitigate biases in healthcare applications [13]. However, existing techniques still struggle with limitations related to imaging depth and resolution, as seen in methods like laser Doppler velocimetry and Doppler optical coherence tomography, underscoring the need for innovative solutions such as PANDA [15].

AI-generated, for reference only.

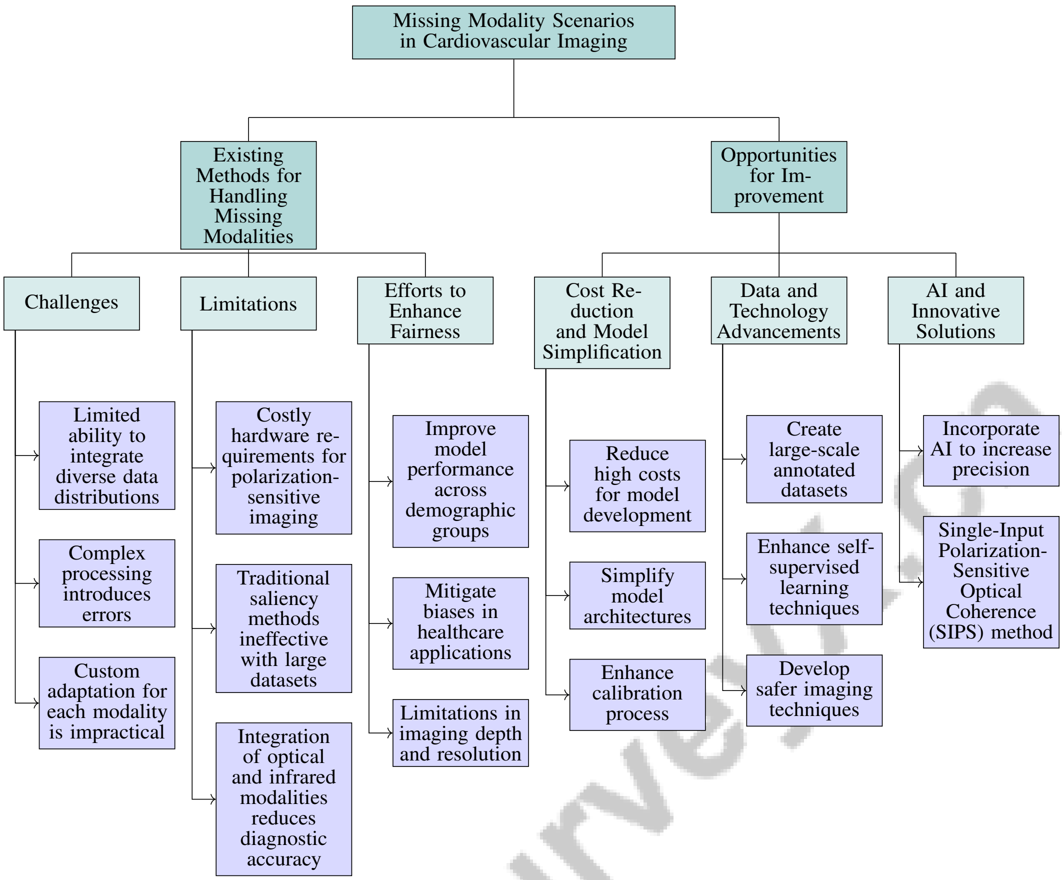


Figure 2: This figure illustrates the challenges and opportunities in handling missing modalities in cardiovascular imaging, highlighting existing method limitations and potential improvements through cost reduction, data advancements, and innovative AI solutions.

3.2 Opportunities for Improvement

The field of cardiovascular imaging offers numerous opportunities for improving methodologies to manage missing modalities. Reducing high costs associated with model development and integration is crucial for broader clinical adoption, ensuring model accuracy and relevance across diverse clinical settings [44]. The complexity of current models, particularly those using multi-parameter calibration based on noisy imaging data, presents additional challenges. Addressing uncertainties in boundary conditions and simplifying model architectures can enhance the calibration process, improving robustness and adaptability to various clinical scenarios [45].

The integration of diverse modalities is further hindered by the lack of large-scale annotated datasets, essential for training robust multimodal learning models. Future research should focus on creating comprehensive datasets and enhancing self-supervised learning techniques to overcome these limitations [27]. Advances in imaging technologies also present significant opportunities. Developing safer imaging techniques and improved image processing methods can enhance diagnostic accuracy. Incorporating artificial intelligence into these processes may further increase the precision and reliability of diagnostic tools [35].

Innovations such as the Single-Input Polarization-Sensitive Optical Coherence (SIPS) method offer promising solutions by simplifying requirements for polarization-sensitive imaging, facilitating broader clinical application [37]. Figure 3 illustrates the key opportunities for improvement in cardiovascular imaging, highlighting challenges in model development, enhancements in data and learning techniques, and technological innovations that can drive advancements in the field.

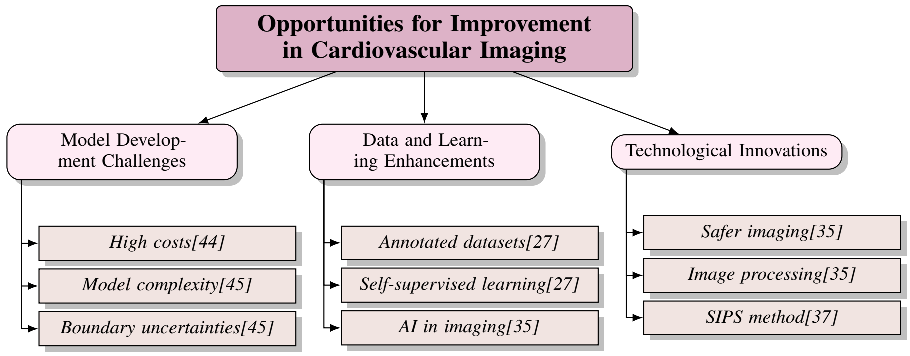


Figure 3: This figure illustrates the key opportunities for improvement in cardiovascular imaging, highlighting challenges in model development, enhancements in data and learning techniques, and technological innovations that can drive advancements in the field.

4 Cross-modal Generation Techniques

4.1 Generative Models for Missing Modality Synthesis

Generative models play a crucial role in synthesizing missing modalities, thereby enhancing the robustness and completeness of multimodal datasets in medical diagnostics. These models leverage advanced machine learning techniques to generate data for absent modalities, significantly boosting diagnostic accuracy and reliability. For example, the Statistical Correlation-Driven Multimodal Fusion (SCM-Fusion) method integrates statistical analysis with human-centered insights to improve pain behavior recognition, demonstrating the efficacy of generative models in addressing missing modalities [33].

In cardiovascular imaging, the PANDA framework exemplifies a generative model that surpasses traditional imaging methods, effectively overcoming challenges posed by missing modalities [15]. This highlights the transformative potential of generative models in enhancing interpretability and accuracy in medical diagnostics.

Feature disentanglement within generative models enables the encoding of multimodal inputs into separate appearance and content codes, which are then combined through a gated fusion process to produce robust representations [7]. This technique is vital for maintaining consistency across modalities and ensuring the reliability of synthesized data.

Innovative architectures like OncoNet, which employ a weakly supervised siamese network, demonstrate the adaptability of generative models by computing differences in tumor burden across sequential imaging exams, thus addressing incomplete datasets [13]. Such applications underscore the capacity of these models to enhance diagnostic processes by compensating for missing modalities.

CSK-Net utilizes contrastive learning-based spectral knowledge distillation alongside a mixed feature exchange strategy, illustrating the adaptability of generative models in improving semantic segmentation across modalities [4]. This adaptability is essential for capturing complex interactions among diverse data sources, ultimately enhancing the diagnostic process.

Despite their promise, generative models face challenges in integrating multimodal data, particularly with respect to data heterogeneity and alignment. Bayoudh et al.'s survey emphasizes the need for effective multimodal representation and fusion techniques, categorizing existing research into six perspectives [26]. This framework provides valuable insights into generating representations in one modality based on another, deepening the understanding of inter-modal relationships.

5 Multimodal Data Fusion

5.1 Data Fusion and Integration Strategies

Data fusion and integration strategies are essential in medical imaging, synthesizing information from multiple modalities to enhance diagnostic accuracy and robustness. These approaches mitigate

AI-generated, for reference only.

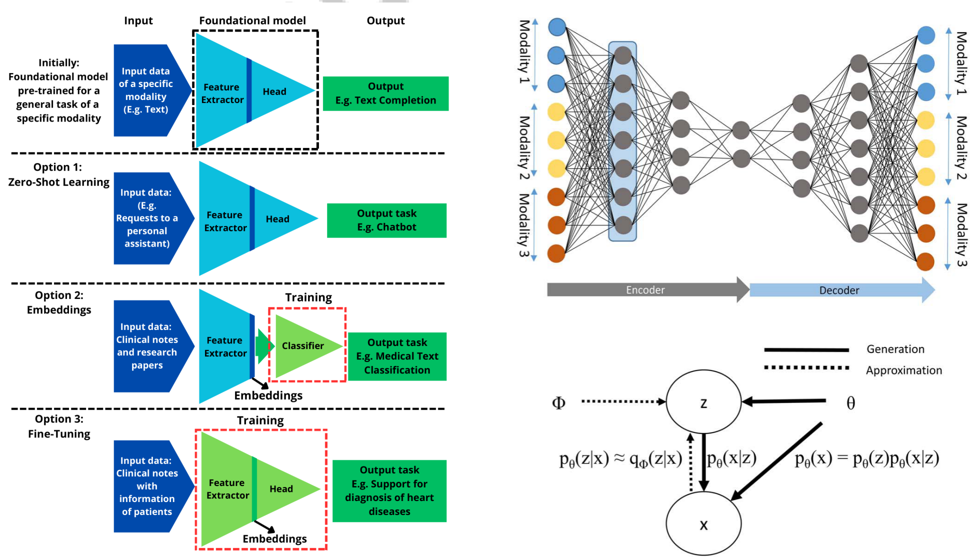
the limitations of single-modality techniques, thereby optimizing diagnostic models. The LoGoCAF architecture exemplifies this by employing a two-branch semantic segmentation framework that learns from hyperspectral and supplementary modalities, effectively capturing local and long-range dependencies [34]. Similarly, the NC2C method leverages a cycleGAN architecture to transform non-contrast CT images into simulated contrast-enhanced images, demonstrating the role of generative models in addressing missing modalities [6]. Action MAE further enhances model robustness by employing self-supervised reconstruction to learn meaningful representations from incomplete data [46].

Weak supervision, as implemented in the OncoNet method, integrates clinical reports to automate treatment response assessments, highlighting the importance of incorporating clinical insights into data fusion frameworks [13]. Such strategies are crucial for improving the scalability and applicability of diagnostic models. By employing advanced multimodal machine learning techniques, data fusion and integration synthesize diverse sources, including electronic health records, imaging, and clinical notes, thereby enhancing the robustness, accuracy, and applicability of diagnostic models [42, 18, 17]. This comprehensive approach reflects clinical experts' decision-making processes and facilitates personalized patient care across various medical domains.

5.2 Advanced Fusion Techniques

Advanced fusion techniques in multimodal data integration are pivotal for addressing the complexities of combining diverse data sources to boost diagnostic precision. The Graph Wavelet Network (GWN) exemplifies this by employing transformer-based attention mechanisms and Long Short-Term Memory (LSTM) to effectively manage and fuse multimodal data, leveraging both temporal and spatial dependencies [47]. The R-BLUE method further illustrates innovation by utilizing rank correlations robust to noise, enhancing diagnostic model integrity amidst noisy and heterogeneous data [9].

Disentangled dense fusion methods and Bayesian generative models address traditional fusion methods' shortcomings by optimizing mutual information and facilitating dense inter-modality feature interaction, reducing computational costs while enhancing data processing reliability across healthcare and high-dimensional datasets [48, 49, 50]. These techniques employ flexible and adaptive models, improving diagnostic processes' robustness and accuracy, paving the way for effective and personalized healthcare solutions.



(a) Three Options for Fine-Tuning a Foundational Model for Different Output Tasks[49]

(b) A Neural Network Model for Multimodal Data[51]

Figure 4: Examples of Advanced Fusion Techniques

As depicted in Figure 4, advanced fusion techniques for multimodal data are illustrated through two approaches. The first, "Three Options for Fine-Tuning a Foundational Model for Different Output Tasks," demonstrates methodologies for adapting a pre-trained foundational model to specific tasks using techniques like zero-shot learning. The second example, "A Neural Network Model for Multimodal Data," features an encoder-decoder framework that integrates multiple modalities into a unified representation, showcasing the model's capability to process complex multimodal data streams [49, 51].

5.3 Explainable AI in Multimodal Fusion

Explainable AI (XAI) is vital for enhancing transparency and interpretability in multimodal data fusion, especially in healthcare, where understanding AI-driven predictions is crucial. XAI provides human-interpretable representations of machine learning models, fostering clinician trust and promoting responsible AI integration into clinical practice. For instance, XAI methods elucidate how different data modalities influence predictions, which is essential in high-stakes environments like healthcare. Systems such as the explainable early warning score (xAI-EWS) maintain predictive performance while clarifying relevant electronic health record data informing their predictions, thereby facilitating safer clinical decision-making [52, 53, 54].

The integration of diverse data types—imaging, genomic, and clinical—requires robust interpretability frameworks to clarify AI models' decision-making processes. XAI techniques deconstruct intricate models, offering insights into how various modalities contribute to diagnostic outcomes. In cardiovascular healthcare, XAI enhances understanding of interactions between different data sources and their impact on patient diagnoses, improving AI systems' reliability and empowering healthcare providers to make informed decisions based on model outputs [39]. This transparency is crucial for addressing concerns related to model bias and fairness, especially in diverse patient populations [11].

Interpretable models utilizing counterfactual explanations enable clinicians to explore alternative scenarios and understand causal relationships between input features and diagnostic outcomes [55]. These approaches enhance model transparency and empower clinicians to validate AI-driven insights, improving clinical decision-making processes. Moreover, integrating XAI into multimodal fusion frameworks helps identify key features and modalities that enhance diagnostic accuracy, ensuring models remain robust even with missing or incomplete data [12].

6 Machine Learning Applications in Cardiovascular Healthcare

6.1 Deep Learning Techniques in Cardiovascular Imaging

Deep learning has revolutionized cardiovascular imaging by extracting intricate patterns, thereby enhancing diagnostic accuracy and clinical decision-making. These techniques enable the integration of multimodal data for comprehensive cardiovascular assessments. The FOAA framework, for instance, outperforms traditional methods in multimodal tumor classification, highlighting the transformative role of advanced algorithms in healthcare diagnostics [5]. Another advancement is the transformation of non-contrast CT images into contrast-enhanced CTAs, potentially minimizing the need for intravenous contrast and its complications [6]. The local-to-global cross-modal attention-aware fusion method further exemplifies the integration of local feature extraction with global feature integration, utilizing convolutional networks and transformer models for improved diagnostic outcomes [34]. Moreover, deep learning applications in augmented reality for ultrasound imaging reflect a shift towards more precise diagnostic tools [1]. Frameworks like CSK-Net demonstrate adaptability to challenging conditions, such as low light, ensuring diagnostic accuracy across diverse clinical scenarios [4].

6.2 Machine Learning for Predictive Analytics

Machine learning (ML) has transformed predictive analytics in cardiovascular healthcare by leveraging diverse datasets and advanced computational techniques. Self-supervised learning enhances multimodal representation learning, crucial for extracting meaningful patterns and improving diagnostic accuracy [27]. The SMIL framework exemplifies ML's robustness in predictive analytics, maintaining accuracy despite missing data [56]. Methods like R-BLUE improve prediction accuracy for non-Gaussian cardiovascular data [9]. Combining meta-learning and adversarial strategies

AI-generated, for reference only.

enhances feature robustness, addressing missing modalities in applications such as brain tumor segmentation [57]. These strategies are vital for developing predictive models that generalize across diverse clinical scenarios. Implementing predictive analytics in online healthcare consultations connects ML applications directly to cardiovascular healthcare, optimizing resource allocation and enhancing patient care [3].

6.3 Interpretable and Robust Machine Learning Models

Integrating interpretable and robust machine learning models is crucial for ensuring AI systems' trustworthiness and effectiveness in healthcare. Interpretable models like EmbraceNet provide transparency in decision-making, fostering clinician trust and safe AI application [58]. This architecture effectively manages cross-modal information, crucial for robustness against data limitations. Action MAE enhances robustness to missing modalities through self-supervised learning, improving model regularization and accuracy [46]. The MMP framework offers robust performance without extensive retraining for missing modalities [10]. These advancements are crucial for managing healthcare data complexities, ensuring consistent diagnostic outcomes. Incorporating fairness metrics in machine learning models promotes equitable healthcare outcomes across diverse demographics. Raza et al. propose a framework enhancing modularity and scalability while fostering fairness [11]. This approach addresses biases, ensuring model reliability. Collaborative efforts among stakeholders are necessary to enhance data management and integration, as highlighted by Maier-Hein et al. Future research should focus on comprehensive data management frameworks and security protocols to protect sensitive patient data. Innovative machine learning techniques adaptable to healthcare challenges, such as poorly labeled disease data and personalized predictions, are vital for leveraging data analytics to improve patient outcomes and optimize resource allocation [59, 42, 20].

7 Clinical Translation and Challenges

7.1 Data Privacy and Security Concerns

Ensuring data privacy and security is paramount in the clinical application of machine learning (ML), especially given the sensitive nature of patient information and the integration of IoT technologies in healthcare. Challenges stem from data complexity, variability, and inconsistent collection and labeling practices, which can jeopardize data integrity and confidentiality [60]. The PANDA framework, supported by the NIH, underscores the need to address these issues in clinical settings [15].

Historical distrust in healthcare systems, exacerbated by non-representative training data, complicates these challenges [61]. Ensuring diverse datasets is essential to mitigate bias and build trust in AI-driven healthcare solutions. Federated learning offers a promising solution by enabling decentralized model training, thus reducing breach risks and enhancing trust [61]. However, it introduces challenges such as increased computational demands and the need for effective arbitration models to handle diverse data sources [60]. High-quality data is crucial, as poor-quality inputs can degrade model performance, highlighting the importance of robust data management strategies [7].

Interoperability among EHR systems remains a significant barrier, necessitating standardized protocols for seamless data integration while safeguarding privacy and security. The quality of incoming data is vital for the adaptability and effectiveness of ML systems, as subpar data quality can severely hinder their performance [60].

7.2 Model Interpretability and Trust

Model interpretability is crucial in clinical settings to foster trust and facilitate the adoption of ML models by healthcare professionals. The complexity of modern algorithms, particularly deep learning models, often results in a black-box nature that complicates clinical integration [52]. This opacity can impede clinicians' understanding and trust in system outputs, limiting their utility in decision-making processes.

Interpretable models can effectively link low-level image features with clinically relevant concepts, enhancing clinician trust by providing outputs that are accurate and meaningful [39]. Additionally, interpretability aids in identifying high-risk individuals through improved prognostic accuracy, enabling proactive healthcare decisions [42]. However, challenges persist in achieving model interpretability

AI-generated, for reference only.

due to the inherent complexity of many ML models, which can obscure decision-making processes and introduce biases in predictions [62]. Current saliency methods often focus on the most probable class, leading to incomplete interpretations that may not align with clinical needs [43].

Research into explainability has provided insights into elucidating ML model behavior, thereby enhancing trust in these systems [54]. However, challenges remain in accurately estimating ranks of common and distinct components in complex models, affecting both performance and interpretability [63].

7.3 Regulatory and Benchmarking Challenges

Benchmark	Size	Domain	Task Format	Metric
PhilHumans[64]	67	Healthcare	Action Anticipation	Accuracy, F1
EHR-SYN[65]	3,184	Cystic Fibrosis	Patient Outcome Classification	AUC-ROC, Accuracy
MVF[66]	1,569	Semantic Segmentation	Semantic Segmentation	mAcc, mIoU
MCA[67]	23,248	Multimodal Sentiment Analysis	Sentiment Recognition	Recall
Zn-13X[68]	7	Materials Science	Data Fusion	F1-score, MRR@10
OCMR[69]	265	Cardiovascular Imaging	Image Reconstruction	PSNR, SSIM

Table 1: This table presents a selection of representative benchmarks used in the evaluation of machine learning models across various domains, including healthcare, semantic segmentation, and materials science. Each benchmark is characterized by its size, domain, task format, and evaluation metric, providing a comprehensive overview of the diverse datasets and assessment criteria employed in the field.

Regulatory and benchmarking challenges are significant barriers to the clinical translation of ML models in cardiovascular healthcare. The complexity and variability of healthcare data necessitate rigorous evaluation frameworks to ensure that these models meet clinical standards and deliver reliable outcomes. Advanced models, such as diffusion models, face limitations due to their computational intensity and long inference times, which can impede practical use in clinical settings [70]. Streamlined processes and efficient algorithms are essential for operating within clinical constraints.

Benchmarking is critical for assessing ML model performance against established clinical criteria. Table 1 provides an overview of various benchmarks that are crucial for evaluating the performance of machine learning models, highlighting the diversity in domain applications and assessment metrics. However, the absence of standardized protocols and evaluation metrics can hinder the integration of AI systems into healthcare workflows. Developing comprehensive benchmarking frameworks is vital for model validation, ensuring AI-driven insights are accurate and clinically relevant. This collaborative effort among researchers, clinicians, and regulatory bodies must create guidelines that address clinical needs while complying with regulatory standards. Such a multidisciplinary approach is essential for navigating the complexities of implementing ML in healthcare, integrating ethical considerations, usability for domain experts, and effective deployment strategies to enhance patient care and safety [38, 18, 71, 72].

Regulatory hurdles are also significant in translating ML models clinically. Compliance with healthcare regulations, particularly regarding patient privacy and data security, is crucial for obtaining regulatory approval and fostering trust among patients and providers. Addressing these regulatory standards is essential for navigating the ethical implications and stakeholder concerns associated with deploying machine learning healthcare applications (ML-HCAs) [59, 20, 17, 71]. The complexity of these regulations necessitates ongoing communication and collaboration among stakeholders to effectively navigate the regulatory landscape.

7.4 Ethical and Fairness Considerations

Ethical and fairness considerations are paramount in deploying machine learning (ML) in healthcare, as these technologies significantly influence clinical decision-making and patient outcomes. The use of ML systems like DeepStroke raises ethical concerns related to biases stemming from the dataset size and diversity used during training [2]. Ensuring that ML models are trained on diverse and representative datasets is vital for avoiding biased outcomes and promoting equitable healthcare delivery.

Aligning multiple modalities in ML applications often presents ethical challenges, particularly when empirical validation in real-world settings is lacking [12]. Insufficient validation can lead to biases in model predictions, adversely affecting patient care. Furthermore, optimizing neural networks for specific imaging sequences can be time-consuming, limiting the scalability of these technologies [16].

Future research should prioritize developing new fusion techniques and improving multimodal dataset quality to address these ethical challenges [26]. By enhancing the robustness and generalizability of ML models, researchers can ensure these technologies yield fair and reliable outcomes across diverse clinical scenarios. Exploring hybrid models that combine various ML approaches may also provide promising solutions to the ethical implications of deploying these technologies in healthcare settings.

7.5 Integration and Interdisciplinary Collaboration

The successful integration of machine learning (ML) into clinical practice relies on robust interdisciplinary collaboration, which is essential for addressing the multifaceted challenges associated with deploying ML in healthcare. Such collaboration is vital for overcoming issues of unidentifiability in ML models, necessitating innovative approaches and the collective expertise of diverse fields to ensure effective clinical integration [73].

Future research should focus on developing ethical frameworks for ML in healthcare, emphasizing algorithmic transparency and the evolving dynamics of the clinician-patient relationship [74]. These frameworks are crucial for guiding the ethical deployment of AI technologies, ensuring alignment with clinical needs and patient expectations.

Exploring causal relationships and advancements in graph neural networks presents significant opportunities for enhancing model interpretability and integration into clinical workflows [62]. By refining interpretability techniques, researchers can develop models that provide clear and actionable insights, facilitating clinician trust and adoption. Addressing challenges posed by missing modalities and developing automatic fusion strategies are critical areas for future research, highlighting the need for interdisciplinary collaboration to optimize ML applications in healthcare [51].

Integrating counterfactual approaches with existing clinical decision support systems offers a promising avenue for real-time diagnostic assistance, further emphasizing the potential of interdisciplinary efforts in advancing healthcare technologies [55]. Establishing standardized evaluation metrics for explainable AI (XAI) methods is vital for assessing their impact on clinical decision-making, ensuring that AI-driven insights are reliable and clinically relevant [54].

The development of such metrics requires collaboration among researchers, clinicians, and policymakers to align AI technologies with clinical standards and regulatory requirements. Existing strengths include developing functional risk prediction models that demonstrate AI's potential to assist in clinical decision-making [38]. Future research should also investigate additional sources of distribution shifts and biases, extending methodologies to handle more complex causal structures [75].

8 Conclusion

Machine learning plays a pivotal role in addressing the challenges posed by missing modality scenarios in cardiovascular healthcare. By integrating diverse medical data, these models enhance diagnostic precision and patient care, despite persistent challenges in clinical implementation due to system compatibility issues. Approaches like MMP demonstrate significant progress, offering improved robustness and performance in scenarios with absent modalities compared to traditional methods.

Advanced models, such as DF-DM, show potential in processing multimodal data with high accuracy and reliability in healthcare contexts. While deep learning techniques hold promise in fields like photoacoustic imaging, further clinical validation is required to ensure real-world applicability beyond simulated environments. The DAPA framework exemplifies advancements in real-time data processing, outperforming conventional methods in speed and accuracy, thus proving its clinical viability.

Multimodal fusion techniques surpass single-modality methods in prediction accuracy, especially in predicting individualized affective experiences. The integration of explainable AI systems, such as xAI-EWS, not only achieves high predictive accuracy but also enhances clinicians' understanding of the underlying decision-making processes, crucial for widespread healthcare adoption. Flex-MoE demonstrates its efficacy in managing various modality combinations, outperforming existing approaches in missing modality scenarios.

Research indicates that thinner neural networks can effectively classify echocardiographic images, facilitating clinical translation in resource-constrained settings. Incorporating demographic factors is essential for improving the fairness and accuracy of diagnostic algorithms. Additionally, standardization in Surgical Data Science is crucial for advancing the field.

Future research should focus on enhancing the explainability of AI systems, refining user interfaces, and examining the long-term effects of AI on clinician-patient interactions. Improved data governance frameworks are necessary to ensure better licensing, persistent access, and comprehensive documentation, thereby enhancing the reliability and fairness of AI models in healthcare. These advancements are vital for fully harnessing the potential of machine learning in cardiovascular healthcare, paving the way for more reliable, equitable, and effective healthcare solutions.

AI-generated, for reference only.

References

[1] Simona Treivase, Alberto Gomez, Jacqueline Matthew, Emily Skelton, Julia A. Schnabel, and Nicolas Toussaint. Screen tracking for clinical translation of live ultrasound image analysis methods, 2020.

[2] Tongan Cai, Haomiao Ni, Mingli Yu, Xiaolei Huang, Kelvin Wong, John Volpi, James Z. Wang, and Stephen T. C. Wong. Deepstroke: An efficient stroke screening framework for emergency rooms with multimodal adversarial deep learning, 2022.

[3] Shuang Geng, Wenli Zhang, Jiaheng Xie, Gemin Liang, Ben Niu, and Sudha Ram. Predicting consultation success in online health platforms using dynamic knowledge networks and multimodal data fusion, 2024.

[4] Aniruddh Sikdar, Jayant Teotia, and Suresh Sundaram. Contrastive learning-based spectral knowledge distillation for multi-modality and missing modality scenarios in semantic segmentation, 2023.

[5] Omnia Alwazzan, Ioannis Patras, and Gregory Slabaugh. Foaa: Flattened outer arithmetic attention for multimodal tumor classification, 2024.

[6] Anirudh Chandrashekar, Ashok Handa, Natesh Shivakumar, Pierfrancesco Lapolla, Vicente Grau, and Regent Lee. A deep learning approach to generate contrast-enhanced computerised tomography angiography without the use of intravenous contrast agents, 2020.

[7] Cheng Chen, Qi Dou, Yueming Jin, Hao Chen, Jing Qin, and Pheng-Ann Heng. Robust multimodal brain tumor segmentation via feature disentanglement and gated fusion. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part III 22*, pages 447–456. Springer, 2019.

[8] Chao Yu, Jiming Liu, Shamim Nemati, and Guosheng Yin. Reinforcement learning in healthcare: A survey. *ACM Computing Surveys (CSUR)*, 55(1):1–36, 2021.

[9] Pengfei Zhang, Gareth W. Peters, Ido Nevat, Keng Boon Teo, and Yixin Wang. Multimodal data fusion of non-gaussian spatial fields in sensor networks, 2019.

[10] Niki Nezakati, Md Kaykobad Reza, Ameya Patil, Mashhour Solh, and M. Salman Asif. Mmp: Towards robust multi-modal learning with masked modality projection, 2024.

[11] Shaina Raza, Parisa Osivand Pour, and Syed Raza Bashir. Fairness in machine learning meets with equity in healthcare, 2023.

[12] Yao Ma, Shilin Zhao, Weixiao Wang, Yaoman Li, and Irwin King. Multimodality in meta-learning: A comprehensive survey. *Knowledge-Based Systems*, 250:108976, 2022.

[13] Anirudh Joshi, Sabri Eyuboglu, Shih-Cheng Huang, Jared Dunnmon, Arjun Soin, Guido Davidzon, Akshay Chaudhari, and Matthew P Lungren. Onconet: Weakly supervised siamese network to automate cancer treatment response assessment between longitudinal fdg pet/ct examinations, 2021.

[14] Cian M Scannell. Automated quantitative analysis of first-pass myocardial perfusion magnetic resonance imaging data, 2021.

[15] Yang Zhang, Joshua Olick-Gibson, Karteekeya Sastry, and Lihong V. Wang. Functional photoacoustic noninvasive doppler angiography in humans, 2024.

[16] Simone Garzia, Patryk Rygiel, Sven Dummer, Filippo Cademartiri, Simona Celi, and Jelmer M. Wolterink. Neural fields for continuous periodic motion estimation in 4d cardiovascular imaging, 2024.

[17] Luis R Soenksen, Yu Ma, Cynthia Zeng, Leonard Boussioux, Kimberly Villalobos Carballo, Liangyuan Na, Holly M Wiberg, Michael L Li, Ignacio Fuentes, and Dimitris Bertsimas. Integrated multimodal artificial intelligence framework for healthcare applications. *NPJ digital medicine*, 5(1):149, 2022.

[18] Adrienne Kline, Hanyin Wang, Yikuan Li, Saya Dennis, Meghan Hutch, Zhenxing Xu, Fei Wang, Feixiong Cheng, and Yuan Luo. Multimodal machine learning in precision health, 2022.

[19] Chaohe Zhang, Xu Chu, Liantao Ma, Yinghao Zhu, Yasha Wang, Jiangtao Wang, and Junfeng Zhao. M3care: Learning with missing modalities in multimodal healthcare data. In *Proceedings of the 28th ACM SIGKDD conference on knowledge discovery and data mining*, pages 2418–2428, 2022.

[20] Marzyeh Ghassemi, Tristan Naumann, Peter Schulam, Andrew L Beam, Irene Y Chen, and Rajesh Ranganath. A review of challenges and opportunities in machine learning for health. *AMIA Summits on Translational Science Proceedings*, 2020:191, 2020.

[21] Farid Ghareh Mohammadi, Farzan Shenavarmasouleh, and Hamid R. Arabnia. Applications of machine learning in healthcare and internet of things (iot): A comprehensive review, 2022.

[22] Jinming Zhao, Ruichen Li, and Qin Jin. Missing modality imagination network for emotion recognition with uncertain missing modalities. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 2608–2618, 2021.

[23] Article.

[24] Fergus Imrie, Stefan Denner, Lucas S. Brunschwig, Klaus Maier-Hein, and Mihaela van der Schaar. Automated ensemble multimodal machine learning for healthcare, 2024.

[25] Shenghua He, Jie Zheng, Akiko Maehara, Gary Mintz, Dalin Tang, Mark Anastasio, and Hua Li. Convolutional neural network based automatic plaque characterization from intracoronary optical coherence tomography images, 2018.

[26] Khaled Bayoudh, Raja Knani, Fayçal Hamdaoui, and Abdellatif Mtibaa. A survey on deep multimodal learning for computer vision: advances, trends, applications, and datasets. *The Visual Computer*, 38(8):2939–2970, 2022.

[27] Naman Goyal. A survey on self supervised learning approaches for improving multimodal representation learning, 2022.

[28] Hu Wang, Yuanhong Chen, Congbo Ma, Jodie Avery, Louise Hull, and Gustavo Carneiro. Multi-modal learning with missing modality via shared-specific feature modelling. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 15878–15887, 2023.

[29] Xiao Song, Xiaodan Zhang, Junzhong Ji, Ying Liu, and Pengxu Wei. Cross-modal contrastive attention model for medical report generation. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 2388–2397, 2022.

[30] Han Qin and Yan Song. Reinforced cross-modal alignment for radiology report generation. In *Findings of the Association for Computational Linguistics: ACL 2022*, pages 448–458, 2022.

[31] Wei Tan, Prayag Tiwari, Hari Mohan Pandey, Catarina Moreira, and Amit Kumar Jaiswal. Multimodal medical image fusion algorithm in the era of big data. *Neural computing and applications*, pages 1–21, 2020.

[32] Yi-Lun Lee, Yi-Hsuan Tsai, Wei-Chen Chiu, and Chen-Yu Lee. Multimodal prompting with missing modalities for visual recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 14943–14952, 2023.

[33] Xingrui Gu, Zhixuan Wang, Irisa Jin, and Zekun Wu. Advancing multimodal data fusion in pain recognition: A strategy leveraging statistical correlation and human-centered perspectives, 2024.

[34] Xuming Zhang, Naoto Yokoya, Xingfa Gu, Qingjiu Tian, and Lorenzo Bruzzone. Local-to-global cross-modal attention-aware fusion for hsi-x semantic segmentation, 2024.

[35] S. K. M Shadekul Islam, MD Abdullah Al Nasim, Ismail Hossain, Md Azim Ullah, Kishor Datta Gupta, and Md Monjur Hossain Bhuiyan. Introduction of medical imaging modalities, 2023.

[36] Lena Maier-Hein, Matthias Eisenmann, Duygu Sarikaya, Keno März, Toby Collins, Anand Malpani, Johannes Fallert, Hubertus Feussner, Stamatia Giannarou, Pietro Mascagni, Hirenkumar Nakawala, Adrian Park, Carla Pugh, Danail Stoyanov, Swaroop S. Vedula, Kevin Cleary, Gabor Fichtinger, Germain Forestier, Bernard Gibaud, Teodor Grantcharov, Makoto Hashizume, Doreen Heckmann-Nötzel, Hannes G. Kenngott, Ron Kikinis, Lars Mündermann, Nassir Navab, Sinan Onogur, Raphael Sznitman, Russell H. Taylor, Minu D. Tizabi, Martin Wagner, Gregory D. Hager, Thomas Neumuth, Nicolas Padoy, Justin Collins, Ines Gockel, Jan Goedeke, Daniel A. Hashimoto, Luc Joyeux, Kyle Lam, Daniel R. Leff, Amin Madani, Hani J. Marcus, Ozanan Meireles, Alexander Seitel, Dogu Teber, Frank Ückert, Beat P. Müller-Stich, Pierre Jannin, and Stefanie Speidel. Surgical data science – from concepts toward clinical translation, 2021.

[37] Georgia L. Jones, Qiaozhou Xiong, Xinyu Liu, Brett E. Bouma, and Martin Villiger. Single-input polarization-sensitive optical coherence tomography through a catheter, 2023.

[38] Roland Roller, Klemens Budde, Aljoscha Burchardt, Peter Dabrock, Sebastian Möller, Bilgin Osmanodja, Simon Ronicke, David Samhammer, and Sven Schmeier. When performance is not enough – a multidisciplinary view on clinical decision support, 2022.

[39] Esther Puyol-Antón, Chen Chen, James R. Clough, Bram Ruijsink, Baldeep S. Sidhu, Justin Gould, Bradley Porter, Mark Elliott, Vishal Mehta, Daniel Rueckert, Christopher A. Rinaldi, and Andrew P. King. Interpretable deep models for cardiac resynchronisation therapy response prediction, 2020.

[40] Amelia Jiménez-Sánchez, Natalia-Rozalia Avlona, Dovile Juodelyte, Théo Sourget, Caroline Vang-Larsen, Anna Rogers, Hubert Dariusz Zając, and Veronika Cheplygina. Copycats: the many lives of a publicly available medical imaging dataset, 2024.

[41] Mihaela Van der Schaar, Ahmed M Alaa, Andres Floto, Alexander Gimson, Stefan Scholtes, Angela Wood, Eoin McKinney, Daniel Jarrett, Pietro Lio, and Ari Ercole. How artificial intelligence and machine learning can help healthcare systems respond to covid-19. *Machine Learning*, 110:1–14, 2021.

[42] Mithun Sarker. Revolutionizing healthcare: the role of machine learning in the health sector. *Journal of Artificial Intelligence General science (JAIGS) ISSN: 3006-4023*, 2(1):36–61, 2024.

[43] Jose Roberto Tello Ayala, Akl C. Fahed, Weiwei Pan, Eugene V. Pomerantsev, Patrick T. Ellinor, Anthony Philippakis, and Finale Doshi-Velez. Signature activation: A sparse signal view for holistic saliency, 2023.

[44] Mark Sendak, Michael Gao, Marshall Nichols, Anthony Lin, and Suresh Balu. Machine learning in health care: a critical appraisal of challenges and opportunities. *EGEMs*, 7(1):1, 2019.

[45] Andreas Mang, Amir Gholami, Christos Davatzikos, and George Biros. Pde-constrained optimization in medical image analysis, 2018.

[46] Sangmin Woo, Sumin Lee, Yeonju Park, Muhammad Adi Nugroho, and Changick Kim. Towards good practices for missing modality robust action recognition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 2776–2784, 2023.

[47] Cong Bao, Zafeirios Fountas, Temitayo Olugbade, and Nadia Bianchi-Berthouze. Multimodal data fusion based on the global workspace theory, 2020.

[48] Chun-An Chou, Xiaoning Jin, Amy Mueller, and Sarah Ostadabbas. Mmdf2018 workshop report, 2018.

[49] David Restrepo, Chenwei Wu, Constanza Vásquez-Venegas, Luis Filipe Nakayama, Leo Anthony Celi, and Diego M López. Df-dm: A foundational process model for multimodal data fusion in the artificial intelligence era, 2024.

[50] Yasin Yilmaz, Mehmet Aktukmak, and Alfred O. Hero. Multimodal data fusion in high-dimensional heterogeneous datasets via generative models, 2021.

[51] Sören Richard Stahlschmidt, Benjamin Ulfenborg, and Jane Synnergren. Multimodal deep learning for biomedical data fusion: a review. *Briefings in Bioinformatics*, 23(2):bbab569, 2022.

[52] Simon Meyer Lauritsen, Mads Kristensen, Mathias Vassard Olsen, Morten Skaarup Larsen, Katrine Meyer Lauritsen, Marianne Johansson Jørgensen, Jeppe Lange, and Bo Thiesson. Explainable artificial intelligence model to predict acute critical illness from electronic health records, 2019.

[53] Mafalda Malafaia, Thalea Schlender, Peter A. N. Bosman, and Tanja Alderliesten. Multifix: An xai-friendly feature inducing approach to building models from multimodal data, 2024.

[54] Yan Jia, John McDermid, Tom Lawton, and Ibrahim Habli. The role of explainability in assuring safety of machine learning in healthcare, 2022.

[55] Jonathan G. Richens, Ciaran M. Lee, and Saurabh Johri. Counterfactual diagnosis, 2020.

[56] Mengmeng Ma, Jian Ren, Long Zhao, Sergey Tulyakov, Cathy Wu, and Xi Peng. Smil: Multimodal learning with severely missing modality. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 2302–2310, 2021.

[57] Aishik Konwer, Xiaoling Hu, Joseph Bae, Xuan Xu, Chao Chen, and Prateek Prasanna. Enhancing modality-agnostic representations via meta-learning for brain tumor segmentation, 2023.

[58] Jun-Ho Choi and Jong-Seok Lee. Embracenet: A robust deep learning architecture for multimodal classification. *Information Fusion*, 51:259–270, 2019.

[59] Munshi Saifuzzaman and Tajkia Nuri Ananna. Towards smart healthcare: Challenges and opportunities in iot and ml, 2024.

[60] Nina Moutonnet, Steven White, Benjamin P Campbell, Saeid Sanei, Toshihisa Tanaka, Hong Ji, Danilo Mandic, and Gregory Scott. Clinical translation of machine learning algorithms for seizure detection in scalp electroencephalography: systematic review, 2024.

[61] Jill A. Kuhlberg, Irene Headen, Ellis A. Ballard, and Donald Martin Jr. au2. Advancing community engaged approaches to identifying structural drivers of racial bias in health diagnostic algorithms, 2023.

[62] Gregor Stiglic, Primoz Kocbek, Nino Fijacko, Marinka Zitnik, Katrien Verbert, and Leona Cilar. Interpretability of machine learning-based prediction models in healthcare. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(5):e1379, 2020.

[63] Ricardo Augusto Borsoi, Konstantin Usevich, David Brie, and Tülay Adali. Personalized coupled tensor decomposition for multimodal data fusion: Uniqueness and algorithms, 2024.

[64] Vadim Liventsev, Vivek Kumar, Allmin Pradhap Singh Susaiyah, Zixiu Wu, Ivan Rodin, Asfand Yaar, Simone Balloccu, Marharyta Beraziuk, Sebastiano Battiato, Giovanni Maria Farinella, Aki Härmä, Rim Helaoui, Milan Petkovic, Diego Reforgiato Recupero, Ehud Reiter, Daniele Riboni, and Raymond Sterling. Philhumans: Benchmarking machine learning for personal health, 2024.

[65] Emily Muller, Xu Zheng, and Jer Hayes. Synthesising electronic health records: Cystic fibrosis patient group, 2022.

[66] Youngjoon Yu, Hong Joo Lee, Byeong Cheon Kim, Jung Uk Kim, and Yong Man Ro. Investigating vulnerability to adversarial examples on multimodal data fusion in deep learning, 2020.

[67] Josiah Bjorgaard. Sparsely multimodal data fusion, 2025.

[68] Calum Green, Sharif Ahmed, Shashidhara Marathe, Liam Perera, Alberto Leonardi, Killian Gmyrek, Daniele Dini, and James Le Houx. Three-dimensional, multimodal synchrotron data for machine learning applications, 2024.

[69] Chong Chen, Yingmin Liu, Philip Schniter, Matthew Tong, Karolina Zareba, Orlando Simonetti, Lee Potter, and Rizwan Ahmad. Ocmr (v1.0)–open-access multi-coil k-space dataset for cardiovascular magnetic resonance imaging, 2020.

[70] Oliver Schad, Julius Frederik Heidenreich, Nils-Christian Petri, Jonas Kleineisel, Simon Sauer, Thorsten Bley, Peter Nordbeck, Bernhard Petritsch, and Tobias Wech. Real-time cardiac cine mri – a comparison of a diffusion probabilistic model with alternative state-of-the-art image reconstruction techniques for undersampled spiral acquisitions, 2024.

[71] Danton S Char, Michael D Abràmoff, and Chris Feudtner. Identifying ethical considerations for machine learning healthcare applications. *The American Journal of Bioethics*, 20(11):7–17, 2020.

[72] Jenna Wiens, Suchi Saria, Mark Sendak, Marzyeh Ghassemi, Vincent X Liu, Finale Doshi-Velez, Kenneth Jung, Katherine Heller, David Kale, Mohammed Saeed, et al. Do no harm: a roadmap for responsible machine learning for health care. *Nature medicine*, 25(9):1337–1340, 2019.

[73] Reza Sameni. Beyond convergence: Identifiability of machine learning and deep learning models, 2023.

[74] Danton S Char, Nigam H Shah, and David Magnus. Implementing machine learning in health care—addressing ethical challenges. *New England Journal of Medicine*, 378(11):981–983, 2018.

[75] Jessica Schrouff, Natalie Harris, Oluwasanmi Koyejo, Ibrahim Alabdulmohsin, Eva Schnider, Krista Opsahl-Ong, Alex Brown, Subhrajit Roy, Diana Mincu, Christina Chen, Awa Dieng, Yuan Liu, Vivek Natarajan, Alan Karthikesalingam, Katherine Heller, Silvia Chiappa, and Alexander D'Amour. Diagnosing failures of fairness transfer across distribution shift in real-world medical settings, 2023.

Disclaimer:

SurveyX is an AI-powered system designed to automate the generation of surveys. While it aims to produce high-quality, coherent, and comprehensive surveys with accurate citations, the final output is derived from the AI's synthesis of pre-processed materials, which may contain limitations or inaccuracies. As such, the generated content should not be used for academic publication or formal submissions and must be independently reviewed and verified. The developers of SurveyX do not assume responsibility for any errors or consequences arising from the use of the generated surveys.