Journal of Al-Qadisiyah for Computer Science and Mathematics Vol.16.(4) 2024,pp.Comp 318–328

Available online at www.qu.edu.iq/journalcm

JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS

ISSN:2521-3504(online) ISSN:2074-0204(print)

A Survey on Exploring Multimodal Fusion in Healthcare: Challenges and Solutions

Riyadh Hussein AL-Mosawi[a*], Ahmed Al-Shammari[b]

[a,b] *Department of Computer Science, College of Computer Science and Information Technology, University of Al-Qadisiyah, Al Diwaniyah, Iraq.*

Email: it.mast.23.5@qu.edu.iq; ahmed.alshammari@qu.edu.iq

ARTICLE INFO

Article history:
Received: 05 /08/2024
Rrevised form: 31 /10/2024
Accepted : 29 /12/2024
Available online: 30 /12/2024

Keywords:

Multimodal Fusion,

Healthcare,

Deep Learning,

Clinical Diagnostics,

Multimodal classification

ABSTRACT

Multimodal fusion in healthcare has emerged as a powerful tool for integrating diverse data sources, including medical images, sensor data, and clinical records, to enhance diagnostic accuracy, treatment planning, and patient monitoring. However, existing surveys need more discussion of the technical details of proposed multimodal fusion approaches and their practical applications. Therefore, we propose a comprehensive survey on using multimodal fusion in healthcare applications. It provides an in-depth examination of the current landscape, discussing key developments that have propelled the field forward as well as ongoing challenges that researchers and practitioners continue to overcome in their efforts to realize the full potential of this technology. The survey explores the fundamentals and core concepts underlying multimodal fusion, explores the methods and techniques that have been developed, and examines the broad applications in various healthcare domains. Moreover, it provides a detailed discussion of the multimodal deep learning classification pipeline, highlighting the critical steps involved in pre-processing, feature extraction, information fusion, and model evaluation. Furthermore, the survey critically analyzes the data integration challenges that remain significant obstacles and identifies promising future directions that hold the potential to shape the future of this rapidly evolving field.

https://doi.org/10.29304/jqcsm.2024.16.41810

1. Introduction

1.1. Overview

Integration of various data modalities has become an important component of high-tech research in the quickly developing field of medical diagnosis and prognosis. Multimodal protrudes as a particularly exciting paradigm, harnessing the special capabilities of various data sources to reach higher predictive performance[1], [2]. To deliver an in-depth and sophisticated analysis of complex and rare diseases, the latest developments in deep learning have enhanced the growth of multimodal techniques. Integration of imaging, genetic, and clinical data has been achieved. Most researchers are attracted to discovering new and effective classification models that use multimodal data to contribute to increasing classification accuracy. This integration has been demonstrated in the healthcare field. The current study focuses on reviewing the latest models for disease prediction and diagnosis to enhance healthcare [3].

*Corresponding author: Riyadh Hussein AL-Mosawi

Email addresses: *it.mast.23.5@qu.edu.iq*

Communicated by 'sub etitor'

One of the most important gains of multimodal classification is its ability to detect complex and unclear relationships between the health and disease axis [1], [4]. These models can identify complex patterns hidden by a single data type by integrating data from diverse sources, such as genetics, health, imaging, and more [2], [5]. This approach also looks very promising, especially in the field of oncology and healthcare. Recent studies have shown that integrating diverse data types has the potential to increase the accuracy of tumor detection [1]. Furthermore, integrating data from diverse sources will help address the problem of data scarcity in some medical fields, opening new paths in healthcare for rare diseases [2], [4].

However, despite the positive results of this multi-modal integration, it has some challenges. The process of harmonizing data that usually comes from different sources, especially in the medical field, constitutes an important and fundamental strategy for developing health models to integrate these different data and to ensure the efficiency, effectiveness and accuracy of these models [2], [5], [6].

1.2. Motivation

Over the past decade, numerous studies have shown significant improvements in diagnostic accuracy by integrating diverse healthcare data sources such as patient imaging, medical records, and genomic data [7]. This is evident from the growing interest and significant growth in studies and research on multimodal medical data classification, as shown in Figure 1. This figure indicates a rapidly advancing promising field with enormous potential to revolutionize healthcare.

Despite these significant advances in AI, especially over the past decade, integrating such diverse data sources still poses significant challenges, including data heterogeneity, privacy, and other computational requirements [39].

This study aims to address these shortcomings by providing an in-depth review of the state-of-the-art multimodal fusion techniques, including input-level, feature-level, and decision-level fusion, and their applications in healthcare. By providing a comprehensive discussion of both the technical foundations and anticipated shortcomings of these approaches, this study aims to guide future research to help healthcare researchers significantly improve diagnostic accuracy and treatment outcomes.

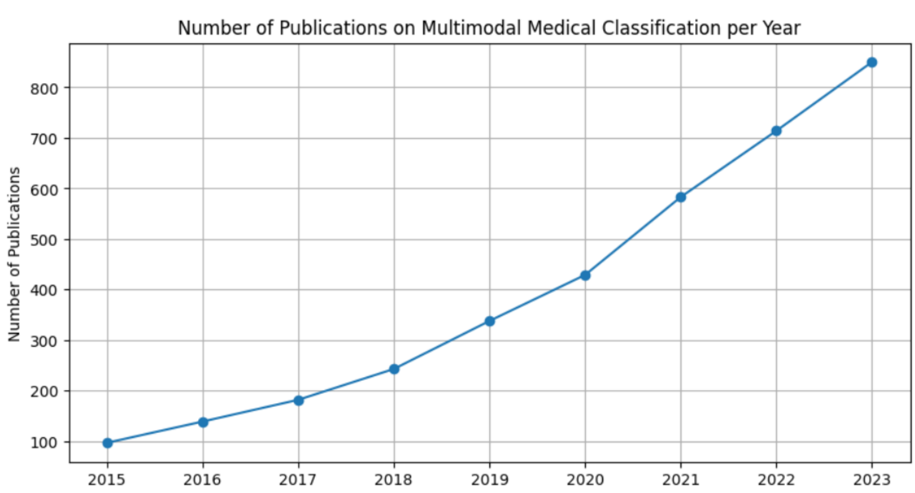

Figure 1 Multimodal medical data classification per year

1.3. Definitions and Terminology

- **Multimodal Fusion**: Mentions the mechanism by which various data sources are combined to improve the analysis and clarification of information. Combining genetic, sensor, electronic health records (EHR), medical imaging, and other types of medical data [7].

- **Deep Learning**: In artificial intelligence, deep learning is a newer extension of machine learning that deals with neural networks that can learn on their own from unlabeled or unstructured data. It has revolutionized medical image processing and has become a powerful and invaluable tool for achieving multimodal fusion. Deep learning has evolved into a useful tool for healthcare and other fields. Due to the amount and proliferation of data, some new challenges have emerged that deep learning has been able to play a major role in addressing [8], [9], [10].

- **Early Developments**: Baseline approaches to data integration in healthcare were predominantly focused on combining simple datasets, primarily structured data from EHRs and basic imaging data[11].

- **Advancements in Technology**: The advent of advanced imaging techniques and the proliferation of wearable technology have contributed extensively to the volume and variety of data available. This has necessitated the development of more sophisticated data fusion techniques[12].

1.4. Key Concepts

- **Data Heterogeneity**: The heterogeneity of data, which includes differences in data format, scale, and temporal alignment of datasets, is one of the main obstacles to multimodal fusion [13].

- **Feature Fusion**: This involves the integration of features extracted from different modalities, which can occur at various levels—early, intermediate, or late—depending on when the integration happens during the data processing pipeline [5].

- **Model Integration**: Discusses the various frameworks and algorithms used to integrate the learning models tailored for each data type, enhancing the decision-making process in clinical settings[1].

2. Methods and Applications of Multimodal Fusion in Healthcare

Traditional fusion methods primarily leverage classical image processing techniques and basic machine learning models to integrate data from multiple sources[4]. These methods can be shown in Figure 2 and categorized as follows:

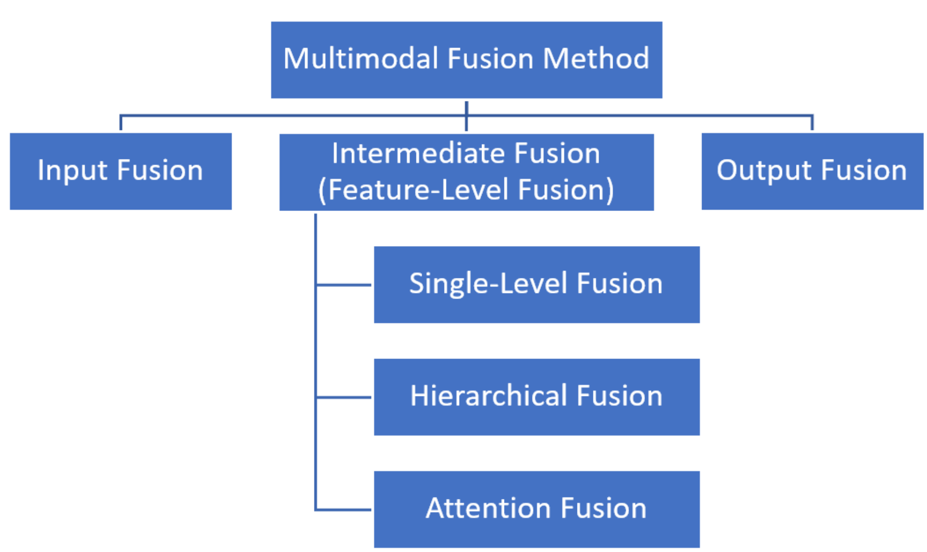

Figure 2 Traditional fusion methods

- **Input Fusion**: Also known as early fusion, this method involves combining raw data from different sources before any processing or analysis is done. Challenges such as spectral degradation or registration errors often accompany input fusion, particularly when merging imaging data from various modalities.

  This technique often requires extensive pre-processing to align and normalize the data to prevent the dominance of one type in the analysis. For example, in neuroimaging studies, MRI and PET scans can be combined at the pixel level to give a comprehensive view of both anatomical and metabolic information, which can aid in the early diagnosis of some neurological diseases. This can be represented as:

  $\mathbf{F_{fused} = \oplus (F_1, F_2, \ldots, F_n)}$ Where:

  - $\mathbf{F_{fused}}$ is the fused input data.
  - $\oplus$ Denotes the fusion operation, which could be concatenation, summation, or another method depending on the data characteristics.
  - $\mathbf{F_i}$ is the feature vector extracted from the **n** modality.
  - **n** is the modality number.

- **Feature-Level Fusion**: This type occurs after the initial feature extraction, where features from different media are combined before entering the final classification or prediction model. This method, which can be implemented through single-level fusion, hierarchical fusion, or attention fusion, has the advantage of preserving the unique information of each pattern but requires many resources and sophisticated techniques to deal with the high dimensions and fluctuations of the feature space.

  This method is based on preserving the information of each pattern and often uses dimensionality reduction techniques to manage the increasing volume of data for processing purposes. The complexity of combining these features from multiple sources is a significant challenge. For example, to diagnose some types of cancer, features from histopathological images can be combined with genetic markers to identify patterns that are more accurate and reliable than those derived from any single pattern. This can be represented as:

  $\mathbf{F_{hierarchical\_level} = \oplus (F_{early\_level}, F_{mid\_level}, \ldots, F_{late\_level})}$ Where:

  - $\mathbf{F_{hierarchical\_level}}$ represents the fused feature vector.
  - $\mathbf{(F_{early\_level}, F_{mid\_level}, \ldots, F_{late\_level})}$ represent feature vectors fused at different levels of the processing pipeline.

- **Decision-Level Fusion**: In this type, separate final decisions are made based on the data decisions derived from each method, and the final decision is reached by combining these multiple decisions together to form the final decision, often using methods such as majority voting, weighted averaging, or more sophisticated decision-making models. This technique is particularly useful in reducing the influence of methods that may perform poorly or be less reliable in some circumstances.

  The benefit of late integration is most apparent when the modalities are very different or when each modality performs well independently but may provide complementary insights when these independent decisions are combined. For example, in inpatient monitoring, decisions from separate analyses such as ECG data and blood biomarkers may be combined to determine a patient's heart health, with each modality providing unique but complementary information about cardiovascular risk. This can be represented as:

  $\mathbf{D_{final} = fusion\ (D_1, D_2, \ldots, D_n)}$ Where:

  - $\mathbf{D_{final}}$ represents the fused feature vector.
  - $\mathbf{D_i}$ represents feature vectors fused at different levels of the processing pipeline.
  - Fusion could be a function like majority voting, averaging, or a learned model combining these decisions.

2.1. Recent Approaches

As deep learning has grown in popularity, sophisticated multimodal fusion techniques have also developed. These techniques are resilient and adaptive, providing opportunities for picture categorization and reliably producing excellent results by managing intricate and extensive data integrations [15], [16], [17]:

- **Deep Learning Fusion**: Utilizes neural networks to perform fusion at various levels. Networks may be designed to perform input, feature, or decision-level fusion, employing architectures that can handle vast amounts of data while learning representations that are more abstract and robust against overfitting.

- **Hybrid Fusion:** Combines traditional and deep learning methods, where initial feature extraction may use traditional algorithms followed by a deep learning model that integrates these features for final analysis. This approach seeks to leverage the strengths of both methodologies, improving performance in complex diagnostic tasks.

- **End-to-End Learning**: A cutting-edge approach where raw data from all modalities are input directly into a deep learning system that autonomously learns to extract features and perform fusion for final decision-making. This method maximizes the potential for uncovering novel insights from data as the model learns which features are most predictive and how best to combine them.

2.2. Applications in Healthcare

Enhancing the accuracy and range of diagnostic imaging relies heavily on multimodal fusion. Clinicians gain a deeper insight into the functional features of the body by combining data from different imaging modalities, such as MRI, CT scans, and PET [18]. This integration contributes to:

- **Advanced Disease Diagnosis**: Improve clarity and dimension for the purpose of helping to diagnose some diseases such as cancer earlier, more accurately and reliably.

- **Precision Healthcare**: Updating and modifying the treatment plan based on increasing information about the characteristics of the disease in each patient.

- **Enhanced Imaging Techniques**: Updating and developing new imaging technologies to collect data from diverse sources in order to provide new and informed insights into complex diseases.

It is worth noting that multimodal fusion has broader applications in healthcare in general. We have expanded our discussion beyond oncology and diagnostics to include cardiology, neurology, and rare disease diagnosis, as shown below. Each field poses unique challenges, and multimodal fusion has shown promising research in improving diagnostic accuracy in most of these areas.

- **Cardiology:** In this area, the integration of diverse data can greatly enhance the diagnosis and monitoring of heart diseases. For example, some recent studies have shown that the integration of ECG data with MRI has clear potential to improve the diagnosis of conditions such as atrial fibrillation. This will greatly contribute to better and faster stroke risk assessment and guide physicians to more accurate, appropriate, and reliable anticoagulant therapies. This is important for improving the care of cardiovascular patients.

- **Neurology:** The benefit of this field lies in the early detection and monitoring of neurological diseases. By integrating data from diverse and different sources for the same patient such as MRI, PET scans, and clinical assessments, healthcare providers can gain a clearer and more comprehensive view of neurological conditions, enhancing early intervention strategies and patient care management.

- **Rare Diseases:** This field often lacks comprehensive data, making diagnosis extremely difficult, as we saw in the previous COVID-19 pandemic. Multimodal integration can bridge this gap by combining genomic data, clinical observations, and diagnostic imaging, facilitating more accurate diagnoses and personalized treatment plans, which are often critical for these patients.

2.3. Treatment Planning and Management

Effective treatment planning is a critical contributor to multimodal integration [19]. This integration allows:

- **Optimize Treatment Protocols:** Using comprehensive, multiple data sets helps design personalized treatment plans that improve outcomes and reduce side effects.
- **Monitor Treatment Efficacy:** Track treatment progress over time by comparing diverse data sources, such as changes in imaging studies correlated with biomarker levels.
- **Support Surgical Interventions**: Enhance surgical precision through advanced visualization techniques that combine real-time imaging data with physiological data during procedures.

2.4. Monitoring and Preventive Medicine

Multimodal fusion can facilitate enhanced patient health status monitoring by merging data from wearables, medical records, and other sources. This can help identify possible health problems early on and enable focused preventive interventions [14]. Additionally, the incorporation of multimedia data can significantly improve healthcare research and development; a few publicly available multimodal datasets are shown in Table 1.

Table 1 - An inventory of multimodal datasets.

Dataset	Body Organ(s)	Medical Diagnosis	EHR
OASIS	Brain	Alzheimer's Disease	Available
SPC	Skin	Skin Lesion	Available
TCGA	rain, Lung, etc	Common Cancer Disease	Available
ABIDE	Brain	Autism Spectrum Disorder	Available
ADHD-200	Brain	Attention Deficit Hyperactivity Disorder	Available
ADNI	Brain	Alzheimer's Disease	Available
TCIA	Brain, Breast, Lung, Kidney, etc.	Common Cancer Disease	Available
PDBP	Brain	Parkinson's disease	Available

2.5. Multimodal Deep Learning classification pipeline

Deep learning combined with multimodal biomedical data fusion is rapidly evolving, with different terminologies creating potential ambiguity. As seen in Figure 3, a five-stage pipeline suggested by [21] offers an organized framework for multimodal medical classification tasks in order to solve this.

- **Pre-processing** is crucial for improving the effectiveness of deep learning. It incorporates methods such as cropping, denoising, picture registration, and feature selection. Data augmentation avoids overfitting and enhances generalization.
- **Deep learning backbone**: The most frequently used pre-trained models were InceptionV3, ResNet50, and VGG16. High-performance network architectures, such as VGG and ResNet, extract high-dimensional features. The development of GPUs, TPUs, and large datasets like ImageNet has propelled deep learning's adaptability across domains[15].
- **Information fusion**: Key for combining data from different modalities. Fusion can be input, intermediate, or output level, employing pooling or complex fusion techniques.
- **Final classifier**: Generates classification results based on multimodal features. Common classifiers include Fully Connected layers, SVM, Random Forest, and Score Merge.

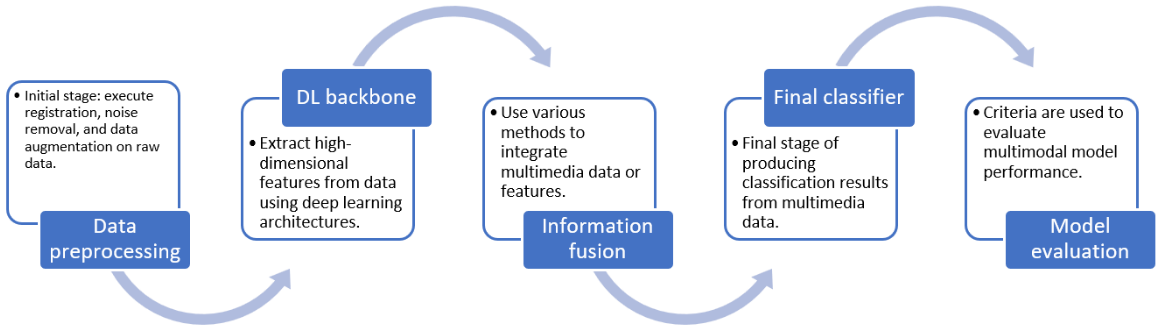

Figure 3 Multimodal classification pipeline

- **Evaluation metrics**: Accuracy, which is the percentage of correct predictions, is the most used measure in classification tasks. Similar to the single-modality tasks, metrics such as accuracy, sensitivity, specificity, precision, F1 score, AUC, and kappa are used to evaluate performance, with a set of results summarized in Table 2.

2.6. Quantitative Analysis of Multimodal Fusion Technique

In this section, we discuss the comparative effectiveness of different multimodal fusion techniques in healthcare applications. Through a quantitative analysis of different fusion techniques, we aim to identify the techniques that provide the most reliable performance metrics across comparable datasets.

Dataset selection: We selected studies with similar characteristics—with a focus on Alzheimer's disease—where each dataset included similar numbers of subjects and used comparable diagnostic methods, ensuring a balanced comparison.

Fusion techniques: We considered including most fusion techniques, such as input fusion, single-level fusion, hierarchical fusion, and output fusion.

Performance metrics: Key metrics such as accuracy, sensitivity, and specificity were considered. These metrics are essential for assessing diagnostic effectiveness in clinical settings.

Tabulated results: Table 2 below provides a clear comparison of different studies using different fusion techniques. This table not only lists the performance metrics, but also puts them in context within a systematic scope where each study and its dataset size are summarized as studies were selected from 2014 to 2023.

Table 2 - A comparison between the outcomes of several fusion techniques.

Research	Year	Fusion Methods	Dataset	NC vs. AD
[17]	2014	Single-level Fusion	398 subjects	Accuracy: 95 %
[18]	2015	Input Fusion	331 subjects	Accuracy: 91 % Sensitivity: 92 % Specificity: 90 %
[19]	2017	Single-level Fusion	202 subjects	Accuracy: 97 % Sensitivity: 95 % Specificity: 98 %
[20]	2019	Single-level Fusion	392 subjects	Accuracy: 98 % Sensitivity: 96 % Specificity: 95 %
[36]	2020	Hierarchical Fusion	500 subjects	Accuracy: 95 % Sensitivity: 94 %

				Specificity: 97 %
[37]	2020	Output Fusion	398 subjects	Accuracy: 99 % Sensitivity: 96 % Specificity: 99 %
[21]	2021	Input Fusion	381 subjects	Accuracy: 94 % Sensitivity: 93 % Specificity: 94 %
[22]	2022	Input Fusion	370 subjects	Accuracy: 93 % Sensitivity: 94 % Specificity: 82 %
[23]	2022	Single-level Fusion	959 subjects	Accuracy: 98 % Sensitivity: 99 % Specificity: 98 %
[38]	2023	Hierarchical Fusion	227 subjects	Accuracy: 97 % Sensitivity: 98 % Specificity: 98 %

3. Challenges and Future Directions

3.1. Data Integration Challenges

Multimodal data fusion faces significant challenges in integrating diverse data types, each with unique formats, scales, and quality[24], [25]:

- **Data Compatibility**: Data compatibility issues arise due to the diverse formats, scales, and sources of healthcare data. One effective solution to this problem is the use of standardized data integration frameworks that can automate the transformation and normalization processes.
  - **Example**: Implementing the Health Level Seven International (HL7) standards for electronic health data facilitates the seamless integration of data from different healthcare systems and devices, ensuring that the fused data is homogenous and interoperable.
- **Data Privacy and Security**: Data privacy remains a major challenge, especially with the increasing use of sensitive medical data [40]. Some advanced encryption methods and privacy-preserving techniques are crucial in addressing these concerns. Below are some solutions that can help address these challenges:
  - **Federated Learning:** This is a powerful and promising approach in which machine learning models are trained across multiple decentralized devices or servers holding local data samples, without sharing them. This method ensures that sensitive data remains on local devices, reducing privacy risks while simultaneously harnessing the collective power of data.
  - **Differential Privacy:** Incorporating differential privacy into data processing ensures that the privacy of individual data entries is maintained when aggregated data is used in machine learning models. This method adds some noise to the data in a way that prevents the identification of any individual's data while not affecting the overall quality of the information.
  - **Homomorphic Encryption:** This technology encrypts data and processes it in its encrypted form, allowing the data to be used in calculations without revealing it. Healthcare organizations can analyze patient data without knowing the encrypted data and use it for research and clinical purposes without compromising privacy.
- **Scalability**: The massive volume of data generated from various forms of healthcare (imaging, genomic data, electronic health records) requires powerful and large computational resources for storage, processing, and

analysis [41]. This increases costs and complicates data management. To address these scaling issues, several solutions can be implemented as follows:

- **Cloud Computing:** Cloud computing technologies provide scalable resources for data storage and computation, allowing healthcare providers to leverage and adjust resources on demand. Cloud services can also facilitate the integration of data across different geographic locations, enhancing research and healthcare.

- **Big Data Technologies:** Big data frameworks, such as Hadoop and Spark, can help manage and process large data sets more efficiently. These technologies are designed to handle massive amounts of data across multiple computing nodes, providing the infrastructure needed to perform large-scale data analytics.

- **Edge Computing:** In this area, data is processed close to the data source rather than in a central data center. This approach allows for reduced latency, reduced bandwidth usage, and improved real-time processing capabilities of health monitoring systems, which is critical for applications such as real-time patient monitoring and telemedicine.

3.2. Technical and Computational Limitations

The depth of multimodal data requires advanced computational techniques that often come with a set of challenges [33], [34]:

- **High Resource Consumption**: The processing power needed to analyze large datasets can be costly.

- **Algorithmic Efficiency**: Creating algorithms capable of managing the fusion of heterogeneous data with high accuracy and speed.

- **Scarcity of Uniform Benchmarks**: The lack of universal standards for evaluating multimodal fusion techniques makes it challenging to evaluate development and contrast methodologies.

4. Future Directions

In view of these issues, Upcoming innovation in the field of multimodal fusion will focus on the following areas:

- **Improved Algorithms**: Develop advanced algorithms that are able to process data complexity, with a focus on minimizing processing demands.

- **Integration of Emerging Technologies**: Exploring the use of evolving technologies such as quantum computing and blockchain to securely and efficiently process and store data.

- **Multidisciplinary Cooperation**: advance innovative solutions spanning data science, medicine, etc.

- **Framework and standard development**: Establishing wide-ranging frameworks and standards for data integration, processing, and analysis to simplify wider implementation and interoperability Standards.

- • **Patient-First Methods**: Focusing on patient-centered designs that prioritize usability, accessibility, and personalized healthcare deliver.

5. Discussion

Studies that used the integration of diverse data have shown high results compared to the results of a single data type, as the accuracy and efficiency of diagnosis for the same disease in complex or rare diseases due to the scarcity of data specific to the type of disease. Furthermore, the development in the field of deep learning has also contributed to a qualitative shift in the diagnosis and classification of diseases. Therefore, the inclusion of deep learning with multimodal data will deal with diversity for the patient data and thus will give greater certainty in the diagnosis and prediction of such diseases in the healthcare sector.

Data integration has a pivotal role in medical diagnosis. For example, the integration of clinical data and patient imaging is necessary in oncology. This has contributed to improving tumor characterization and prognosis predictions. This approach not only enhances the precision of diagnostics but also contributes significantly to the integration of advanced imaging techniques and modern healthcare.

Despite the promising developments mentioned above in the field of multimedia data, this study presents some challenges facing the integration of diverse data. The issues of integrating diverse data, in addition to the huge computational costs that require high computational resources and the complexity of the algorithm, pose significant obstacles to the effective integration of multimedia data. Looking forward, the researchers suggest developing more algorithms that are capable of dealing with the diversity of data sources. The possibility of integrating emerging technologies such as quantum computing may also contribute to reducing this complexity. It will provide greater and faster opportunities to deal with diverse information, especially in the field of healthcare, making it a very important and promising area of research in the coming years.

6. Conclusion

Our survey underscores the pivotal and promising role of multimodal fusion technologies in revolutionizing healthcare, particularly by enhancing the accuracy of disease diagnosis. We explored innovative models that skillfully manage and integrate diverse and disparate data sources, including genomic data, medical imaging, and electronic health records. These models have also proven to outperform traditional approaches not only by improving diagnostic accuracy but also by uncovering complex patterns in some rare diseases and providing predictive insights. Such developments contribute to the provision of broad healthcare solutions that leverage the unique strengths of each type of data being captured.

The great advances in the field of deep learning have stimulated the development of advanced algorithms capable of dealing with the complexities associated with diverse data. This paves the way for promising, more reliable, and accurate medical practices and motivates researchers to develop personalized treatment strategies that promise greater effectiveness.

However, the process of integrating multimodal data into healthcare presents several challenges, including data compatibility and integration issues and the large computational resources required. This requires significant investments in technological infrastructure and improvements in computational methods to improve the operational efficiency and speed of these advanced models.

However, the potential of multimodal integration in healthcare is still very promising. It continues to inspire many researchers to enter this field, which requires collaborative efforts across disciplines to overcome the current limitations. Continuous research and technological advancements in this field are greatly contributing to exploiting the great potential offered by multimodal integration, ultimately paving the way for more accurate, efficient, and personalized healthcare solutions.

References

[1] A. W. A. N. D. A. T. A. N. D. R. P. R. A. N. D. P. S. A. N. D. G. Rasool, “Multimodal Data Integration for Oncology in the Era of Deep Neural Networks: A Review,” Jul. 2023. doi: 10.48550/arXiv.2303.

[2] E. Warner et al., “Multimodal Machine Learning in Image-Based and Clinical Biomedicine: Survey and Prospects,” Int J Comput Vis, Apr. 2024, doi: 10.1007/s11263-024-02032-8.

[3] Z. Ch. Oleiwi, E. N. AlShemmary, and S. Al-augby, “Adaptive Features Selection Technique for Efficient Heart Disease Prediction,” Journal of Al-Qadisiyah for Computer Science and Mathematics, vol. 15, no. 1, Feb. 2023, doi: 10.29304/jqcm.2023.15.1.1137.

[4] A. Kline et al., “Multimodal machine learning in precision health: A scoping review,” NPJ Digit Med, vol. 5, no. 1, p. 171, 2022.

[5] C. C. A. N. D. J. W. A. N. D. Y. W. A. N. D. S. Z. A. N. D. Z. A. A. N. D. L. A. C. A. N. D. K. T. W. A. N. D. B. A. L. A. N. D. Y. Huo, “Deep Multi-modal Fusion of Image and Non-image Data in Disease Diagnosis and Prognosis: A Review,” Jul. 2022. doi: 10.48550/arXiv.2203.

[6] M. Yaseen Abdullah, “Optimization Techniques for Human Multi-Biometric Recognition System,” Journal of Al-Qadisiyah for Computer Science and Mathematics, vol. 15, no. 3, Oct. 2023, doi: 10.29304/jqcsm.2023.15.31292.

[7] M. A. Al-Antari, “Artificial Intelligence for Medical Diagnostics—Existing and Future AI Technology!,” Diagnostics, vol. 13, no. 4, p. 688, Feb. 2023, doi: 10.3390/diagnostics13040688.

[8] F. Mohsen, H. Ali, N. El Hajj, and Z. Shah, “Artificial intelligence-based methods for fusion of electronic health records and imaging data,” Sci Rep, vol. 12, no. 1, p. 17981, Oct. 2022, doi: 10.1038/s41598-022-22514-4.

[9] S.-C. Huang, A. Pareek, S. Seyyedi, I. Banerjee, and M. P. Lungren, “Fusion of medical imaging and electronic health records using deep learning: a systematic review and implementation guidelines,” NPJ Digit Med, vol. 3, no. 1, p. 136, 2020.

[10] R. J. Al Gharrawi and A. A. Al-Joda, “A Survey of Medical Image Analysis Based on Machine Learning Techniques,” Journal of Al-Qadisiyah for Computer Science and Mathematics, vol. 15, no. 1, Feb. 2023, doi: 10.29304/jqcm.2023.15.1.1139.

[11] R. N. Sadoon and A. M. Chaid, “Deep Learning Technique for The Classification of Lung Diseases from X-ray Images,” Journal of Al-Qadisiyah for Computer Science and Mathematics, vol. 16, no. 2, Jun. 2024, doi: 10.29304/jqcsm.2024.16.21544.

[12] A. B. Menegotto, C. D. L. Becker, and S. C. Cazella, “Computer-aided diagnosis of hepatocellular carcinoma fusing imaging and structured health data,” Health Inf Sci Syst, vol. 9, no. 1, p. 20, 2021, doi: 10.1007/s13755-021-00151-x.

[13] G. Muhammad, F. Alshehri, F. Karray, A. El Saddik, M. Alsulaiman, and T. H. Falk, “A comprehensive survey on multimodal medical signals fusion for smart healthcare systems,” Information Fusion, vol. 76, pp. 355–375, 2021, doi: https://doi.org/10.1016/j.inffus.2021.06.007.

[14] E. Warner et al., “Multimodal Machine Learning in Image-Based and Clinical Biomedicine: Survey and Prospects,” Int J Comput Vis, 2024, doi: 10.1007/s11263-024-02032-8.

[15] A. Holzinger, B. Malle, A. Saranti, and B. Pfeifer, “Towards multi-modal capability with Graph Neural Networks enabling information fusion for explainable AI,” Information Fusion, vol. 71, pp. 28–37, 2021, doi: https://doi.org/10.1016/j.inffus.2021.01.008.

[16] Z. Salahuddin, H. C. Woodruff, A. Chatterjee, and P. Lambin, “Transparency of deep neural networks for medical image analysis: A review of interpretability methods,” Comput Biol Med, vol. 140, p. 105111, 2022.

[17] E. H. Abdulsaed, M. Alabbas, and R. S. Khudeyer, “Optimizing the Architecture of Convolutional Neural Networks Using Modified Salp Swarm Algorithm,” Journal of Al-Qadisiyah for Computer Science and Mathematics, vol. 16, no. 1, Mar. 2024, doi: 10.29304/jqcsm.2024.16.11450.

[18] K. Doi, “Diagnostic imaging over the last 50 years: research and development in medical imaging science and technology,” Phys Med Biol, vol. 51, no. 13, p. R5, 2006, doi: 10.1088/0031-9155/51/13/R02.

[19] P. Shah et al., “Artificial intelligence and machine learning in clinical development: a translational perspective,” NPJ Digit Med, vol. 2, no. 1, p. 69, Jul. 2019, doi: 10.1038/s41746-019-0148-3.

[20] O. Amft, “How Wearable Computing Is Shaping Digital Health,” Jul. 2018. doi: 10.1109/mprv.2018.011591067.

[21] W. C. Sleeman IV, R. Kapoor, and P. Ghosh, “Multimodal classification: Current landscape, taxonomy and future directions,” ACM Comput Surv, vol. 55, no. 7, pp. 1–31, 2022.

[22] S. Hameed Al-Amiry, “Survey on Diagnosing Retina Diseases in Optical Coherence Tomography Images Based on Deep Learning,” Journal of Al-Qadisiyah for Computer Science and Mathematics, vol. 16, no. 2, Jun. 2024, doi: 10.29304/jqcsm.2024.16.21568.

[23] H.-I. Suk, S.-W. Lee, D. Shen, and A. D. N. Initiative, “Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis,” Neuroimage, vol. 101, pp. 569–582, 2014.

[24] S. Liu et al., “Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer’s disease,” IEEE Trans Biomed Eng, vol. 62, no. 4, pp. 1132–1140, 2014.

[25] J. Shi, X. Zheng, Y. Li, Q. Zhang, and S. Ying, “Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer’s disease,” IEEE J Biomed Health Inform, vol. 22, no. 1, pp. 173–183, 2017.

[26] F. Zhang, Z. Li, B. Zhang, H. Du, B. Wang, and X. Zhang, “Multi-modal deep learning model for auxiliary diagnosis of Alzheimer’s disease,” Neurocomputing, vol. 361, pp. 185–195, 2019.

[27] J. Song, J. Zheng, P. Li, X. Lu, G. Zhu, and P. Shen, “An effective multimodal image fusion method using MRI and PET for Alzheimer’s disease diagnosis,” Front Digit Health, vol. 3, p. 637386, 2021.

[28] Z. Kong, M. Zhang, W. Zhu, Y. Yi, T. Wang, and B. Zhang, “Multi-modal data Alzheimer’s disease detection based on 3D convolution,” Biomed Signal Process Control, vol. 75, p. 103565, 2022.

[29] M. Abdelaziz, T. Wang, and A. Elazab, “Fusing multimodal and anatomical volumes of interest features using convolutional auto-encoder and convolutional neural networks for Alzheimer's disease diagnosis,” Front Aging Neurosci, vol. 14, p. 812870, 2022.

[30] Y. A. Mohammed and E. G. Saleh, “Investigating the Applicability of Logistic Regression and Artificial Neural Networks in Predicting Breast Cancer,” Journal of Al-Qadisiyah for Computer Science and Mathematics, vol. 12, no. 2, Jul. 2020, doi: 10.29304/jqcm.2020.12.2.697.

[31] C. Cui et al., “Deep multimodal fusion of image and non-image data in disease diagnosis and prognosis: a review,” Progress in Biomedical Engineering, vol. 5, no. 2, p. 022001, 2023, doi: 10.1088/2516-1091/acc2fe.

[32] R. M. A. N. D. F. W. A. N. D. S. W. A. N. D. X. J. A. N. D. J. T. Dudley, “Deep learning for healthcare: review, opportunities and challenges,” Jul. 2017. doi: 10.1093/bib/bbx044.

[33] M. Salvi et al., “Automated assessment of glomerulosclerosis and tubular atrophy using deep learning,” Computerized Medical Imaging and Graphics, vol. 90, p. 101930, 2021, doi: https://doi.org/10.1016/j.compmedimag.2021.101930.

[34] A. Abernethy et al., “The Promise of Digital Health: Then, Now, and the Future,” NAM Perspectives, vol. 6, no. 22, Jun. 2022, doi: 10.31478/202206e.

[35] S. R. Stahlschmidt, B. Ulfenborg, and J. Synnergren, “Multimodal deep learning for biomedical data fusion: a review,” Brief Bioinform, vol. 23, no. 2, Mar. 2022, doi: 10.1093/bib/bbab569.

[36] Zhang, T., & Shi, M. (2020). Multi-modal neuroimaging feature fusion for diagnosis of Alzheimer’s disease. Journal of neuroscience methods, 341, 108795.

[37] Fang, X., Liu, Z., & Xu, M. (2020). Ensemble of deep convolutional neural networks based multimodality images for Alzheimer's disease diagnosis. IET Image Processing, 14(2), 318-326.

[38] Chen, H., Guo, H., Xing, L., Chen, D., Yuan, T., Zhang, Y., & Zhang, X. (2023). Multimodal predictive classification of Alzheimer's disease based on attention-combined fusion network: Integrated neuroimaging modalities and medical examination data. IET Image Processing, 17(11), 3153-3164.

[39] Sun, Z., Wang, Z., & Xu, Y. (2023). Privacy protection in cross-platform recommender systems: techniques and challenges. Wireless Networks, 1-10.

[40] Rahman, M. A. (2023, December). A Survey on Security and Privacy of Multimodal LLMs-Connected Healthcare Perspective. In 2023 IEEE Globecom Workshops (GC Wkshps) (pp. 1807-1812). IEEE.

[41] Mohsen, F., Ali, H., El Hajj, N., & Shah, Z. (2022). Artificial intelligence-based methods for fusion of electronic health records and imaging data. Scientific Reports, 12(1), 17981.