Review on multi-modal AI models to integrate imaging and omics data

Matteo Di Vincenzo[1*] and Wilson Silva[2]

Utrecht University, AI for life sciences group, Netherlands.

*Corresponding author(s). E-mail(s): m.divincenzo@students.uu.nl;
Contributing authors: w.j.dossantossilva@uu.nl;

Abstract

Multimodal artificial intelligence (AI) revolutionizes biomedical research by integrating electronic health records and imaging data. This review succinctly explores the imperative of multimodal AI, outlines covered topics, examines existing types, discusses limitations, and proposes future research.

In cancer research, multimodal AI is used to predict survival rates and identify molecular subtypes, exemplified by methods like Cancer Integration via Multikernel Learning (CIMLR). Extending beyond cancer, it contributes to various facets of precision medicine and radiation therapy. Precision medicine benefits from AI-driven radiomics, a cost-effective method for biomarker identification through medical image analysis. The adaptability of multimodal AI is evident in its application to single-cell analysis, showcasing its potential across diverse modalities. However, managing extensive datasets combining imaging and omics data poses a challenge. The focus shifts to refining representations and effectively combining modalities. Statistical and machine learning approaches, including supervised classification algorithms, regression analysis, unsupervised clustering, and network analysis, play pivotal roles. Deep learning techniques, particularly Convolutional Neural Networks (CNNs), succeed in image recognition and genomics analysis. Models like VGG19-CNN and self-normalization networks integrate image and omics data, enhancing predictions related to cancer-specific survival and subtype classification. In summary, multimodal AI in biomedical research holds promise for disease diagnosis, subtyping, and precision medicine. This work provides a concise overview of recent progress and potential, emphasizing its transformative role. The synergy of electronic health records and imaging data facilitated by AI enhances our understanding of biological processes, heralding a new era in personalized medicine.

1 Layman Summary

In recent years, scientists have made significant strides in combining different types of data using advanced computer models. This exciting field, known as multimodal artificial intelligence (AI), involves merging electronic health records and imaging data. This integration has proven particularly valuable in biological and medical studies. Essentially, researchers are using AI to blend information from medical images with genetic and molecular data to gain a more complete understanding of how our bodies work. A major focus of this approach is on cancer research. Scientists are using multimodal AI to predict survival rates in liver cancer and identify specific types of ovarian cancer. One method, called Cancer Integration via Multikernel Learning (CIMLR), uses AI to integrate different types of data and reveal molecular subtypes of cancer. This kind of AI is transforming our understanding of cancer and could lead to new and improved ways of treating the disease. But the impact of multimodal AI isn't limited to cancer research. playing a pivotal role in various aspects of precision medicine and radiation therapy. In precision medicine, scientists are using AI to analyze medical images and identify biomarkers—indicators that help tailor treatments to individual patients. This process is non-invasive, quick, and cost-effective. Moreover, when it comes to studying single cells in the body, integrating data from different sources using multimodal AI is proving to be a powerful tool. The combination of imaging and omics data through multimodal AI is opening up exciting possibilities in the medical field. This approach has the potential to revolutionize how we diagnose diseases, categorize them, and tailor treatments to each patient. Cutting-edge computer models and AI techniques are at the heart of this transformative approach, helping scientists unravel the complexities of diseases at a microscopic level. To understand how multimodal AI works, we need to look at the progress made in medical imaging technologies like CT scans, MRIs, and PET scans. These imaging techniques provide detailed pictures of the inside of our bodies, helping doctors see and understand anatomical structures and processes. On the other side, omics data, which includes information about genes, proteins, and other molecules, provides a comprehensive view of an organism's genetic and molecular makeup. However, managing the vast amount of data generated by combining imaging and omics data is a major challenge. This involves integrating information from various sources such as genetic data, information about proteins, and medical images. This integration is essential for conducting meaningful analyses, addressing challenges like selecting important features, classifying different aspects, and understanding relationships between different pieces of information. To make sense of this complex data, scientists use a variety of statistical and machine learning approaches. These encompass supervised classification algorithms, which facilitate the identification of patterns in the data, and unsupervised clustering, which organizes similar data points into groups. Deep learning techniques, including Convolutional Neural Networks (CNNs), play a pivotal role in deciphering complex biomedical data. Notably, some studies stand out for utilizing deep learning, especially CNNs, to integrate image and omics data, showcasing the versatility of these models across both supervised and unsupervised learning paradigms. Models like VGG19-CNN and self-normalization networks are improving predictions related to cancer survival and subtypes. One model, called PAGE-NET, uses a patch-wise texture-based CNN to integrate genomic data, patient

age, and histological images. Multimodal neural networks, like MM-Net, are even predicting how patients will respond to drugs by combining drug information, gene expression data, and histological features.
In summary, the integration of multimodal AI in medicine holds great promise for improving disease diagnosis, classification, and treatment. This overview provides insights into the recent progress, challenges, and future possibilities of multimodal AI, showcasing its crucial role in advancing medical research and healthcare.

2 Introduction

The integration of multimodal artificial intelligence (AI) in the medical field has garnered substantial attention due to its transformative potential in revolutionizing disease diagnosis and treatment. This paradigm shift is particularly evident in the convergence of imaging and omics data, which has opened new avenues for understanding diseases at a molecular level. As the volume of omics data continues to accumulate, the prospect of developing computational models that seamlessly integrate multimodal data sources becomes increasingly feasible [1]. This has spurred the development of advanced AI-based methods for the fusion of electronic health records and imaging data, and these methods are progressively finding applications in both biological and medical studies [2].
One significant approach to integrating images with multi-omics data involves deriving phenotypic information from imaging data. This information is then utilized as annotations to enhance the interpretation of omics data, contributing to a more holistic understanding of the biological landscape [3]. Furthermore, the imperative for multimodal learning paradigms has been underscored, emphasizing their role in providing reliable diagnoses of diseases such as cancer by combining omics, bioimaging, and clinical outcomes [4].
The potential of multimodal AI in the medical field is particularly conspicuous in cancer research. Studies have not only highlighted the lack of concerted efforts in integrating multi-omics data for predicting survival in liver cancer across multiple patient cohorts [5], but they have also showcased successful integration of medical images with multi-omics analysis in ovarian cancer using artificial intelligence [6]. Moreover, innovative approaches like Cancer Integration via Multikernel Learning (CIMLR) have demonstrated the integration of multi-omic data to unveil molecular subtypes of cancer, further underlining the potential of multimodal AI in advancing cancer research [7]. Beyond cancer, the integration of imaging and omics data has shown promise in precision medicine and radiation therapy. Radiomics, coupled with artificial intelligence, has emerged as a novel approach for precision medicine in radiation therapy, leveraging multimodal medical images for non-invasive, fast, and cost-effective biomarker identification [8]. Additionally, computational methods for the analysis and integration of single-cell omics data across different modalities have been summarized, highlighting the potential of multimodal AI in advancing single-cell analysis [9].
In conclusion, the integration of multimodal AI for combining imaging and omics data in the medical field presents a groundbreaking opportunity to advance disease diagnosis, subtyping, and precision medicine. The convergence of imaging and omics

data, coupled with advanced computational models and AI techniques, offers a transformative approach to understanding and addressing complex diseases at a molecular level. This comprehensive review aims to provide a thorough overview of the current state of multimodal AI integration in the medical field, offering insights into the latest advancements, challenges, and future prospects in this rapidly evolving domain.

3 Background

Artificial intelligence (AI) that integrates imaging data and omics data has ushered in a transformative era in biomedical research and healthcare. This section aims to elucidate the fundamental principles of medical imaging, pathology, genomics, proteomics, computer vision, and the integration of machine learning and deep learning techniques for a comprehensive understanding of this scientifc review.

3.1 Imaging Data

In the realm of medical imaging, various techniques play crucial roles in providing comprehensive insights into anatomical and physiological aspects. Computed Tomography (CT), a non-invasive imaging technique utilizing X-rays, generates cross-sectional images of anatomical structures. Through the application of mathematical algorithms, CT reconstructs detailed images, allowing for the precise visualization of internal tissues and organs. Magnetic Resonance Imaging (MRI) works by employing strong magnetic fields and radiofrequency pulses. This technique creates detailed images of soft tissues, leveraging variations in water content and molecular properties to excel in capturing high-resolution anatomical details. Positron Emission Tomography (PET) introduces a different dimension by involving the injection of radiolabeled tracers to assess metabolic activity in tissues. PET detects positron emissions resulting from tracer decay, providing functional information about physiological processes at the molecular level. The synergy of PET and CT in PET/CT integration combines anatomical and functional information. While CT offers detailed structural images, PET reveals metabolic activity, enabling a comprehensive assessment of pathology and disease progression.

Turning our attention to histological techniques, Hematoxylin and Eosin (H&E) staining emerges as a standard practice. This technique employs hematoxylin to stain cell nuclei blue and eosin to stain cytoplasm and extracellular structures pink. The staining method facilitates microscopic examination and the characterization of tissue morphology. Ultimately, the emergence of Whole Slide Imaging (WSI) introduces digitization to the field of histology. This technique involves digitizing entire histological slides to create high-resolution digital images. This facilitates computational analysis, remote access, and the application of machine learning algorithms for automated pathology.

In the domain of medical image analysis, computer vision techniques are applied to interpret and diagnose medical images, encompassing the segmentation of anatomical structures, detection of abnormalities, and quantitative analysis of imaging features. Computer vision proves invaluable in Histopathology Image Analysis, as machine learning algorithms excel in identifying patterns in digitized pathology slides, aiding pathologists in diagnosing diseases, quantifying biomarkers, and predicting patient

outcomes. Furthermore, the field of Radiomics involves the extraction and analysis of quantitative features from medical images. Prediction models in Radiomics can rely on predefined functions or learned features, often employing techniques such as Deep Learning. Computer vision plays a pivotal role in deciphering complex patterns in radiological images, facilitating the identification of imaging biomarkers for disease characterization.

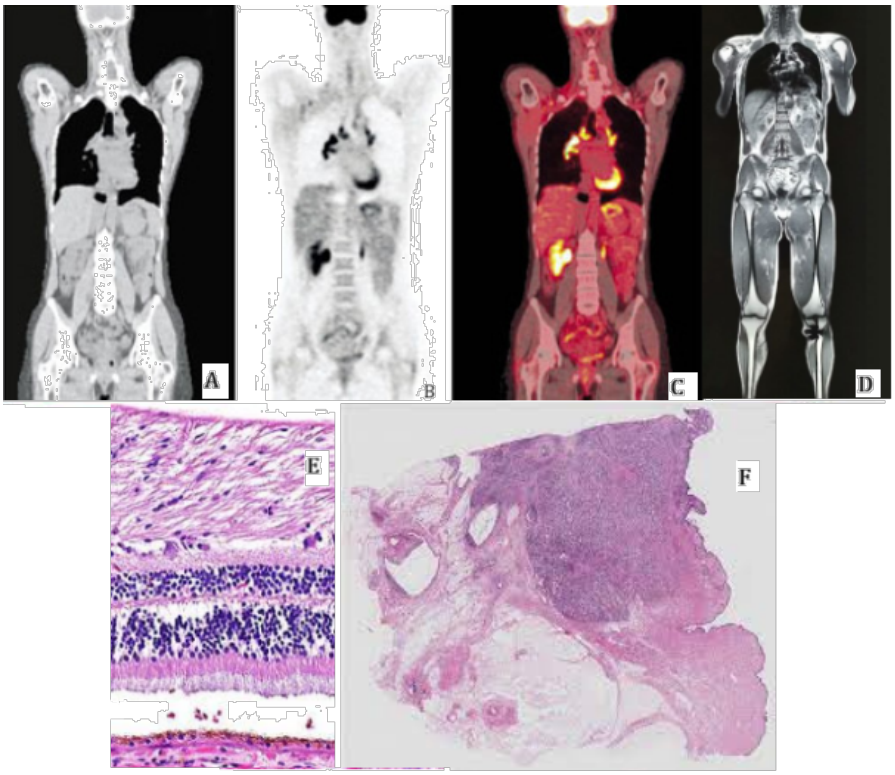

Fig. 1 Comparison of different imaging technologies: a:CT[10], b: PET[10], d: PET/CT[10], e: H&E[11], f:WSI[12]

3.2 Omics Data

In the comprehensive exploration of molecular biology, various branches come together to contribute significantly to our understanding of genetic and molecular processes, culminating in what we describe as omics.

Genomics involves a thorough exploration of an organism's complete genomic content, encompassing genes, non-coding regions, and structural variations. High-throughput sequencing technologies facilitate the analysis of DNA sequences, aiding

in the identification of genetic variations and functional elements. Moving to **Epigenomics**, this field delves into modifications to DNA and histone proteins regulating gene expression without altering the underlying DNA sequence. Techniques such as chromatin immunoprecipitation sequencing (ChIP-seq) unveil intricate epigenetic landscapes. Advancing to **Transcriptomics**, the focus is on the study of RNA transcripts produced in a specific biological context. Leveraging RNA sequencing (RNA-Seq), we can quantify and characterize the transcriptome, gaining insights into gene expression patterns and alternative splicing events. In the realm of **Proteomics**, a systematic study of an organism's complete complement of proteins takes place. Employing mass spectrometry-based techniques, such as liquid chromatography-mass spectrometry (LC-MS), enables the identification and quantification of proteins, unraveling complex cellular processes.
While both **RNA-Seq** and **RNA-microarrays** significantly contribute to molecular insights, RNA-Seq, is a high-throughput sequencing method that, quantifies RNA transcripts, providing a detailed and unbiased view of the transcriptome. Whereas, RNA-microarrays use complementary DNA (cDNA) probes to measure gene expression levels, allowing the simultaneous analysis of multiple genes.

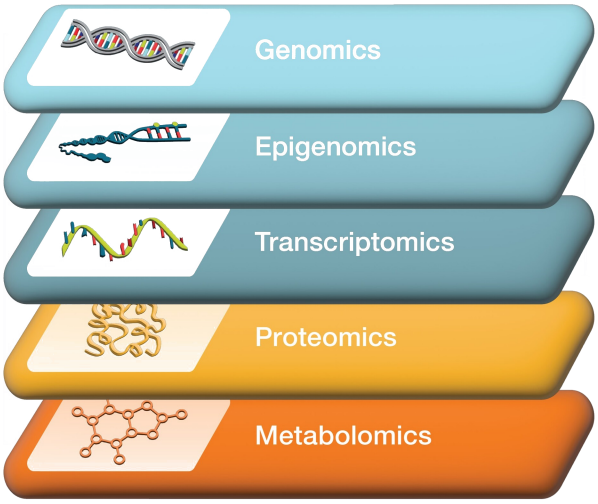

Fig. 2 Integration of diverse Omics data, demonstrating their collective contribution to a holistic analysis.[13]

4 Data

4.1 Multimodal Data

Contemporary biomedical research faces a significant challenge in effectively managing features derived from extensive datasets that combine imaging and high-throughput omics data. In this article, we refer to this challenge as 'multimodal data', representing a collection of information from diverse features and sample sets. This data, generated from heterogeneous sources, provides complementary information crucial for characterizing biological samples, events, or systems.
In the context of a study, multimodal data processing entails the integration of information from a minimum of two among the following categories: anatomical pathology, genomics, epigenomics, transcriptomics, proteomics, and medical imaging data. While our focus revolves around these specific data categories, it's important to note that multimodal AI extends beyond these domains, encompassing a broader spectrum of data modalities. The integration of multimodal data is pivotal for conducting data-driven analyses, addressing challenges such as feature selection, classification, regression, unsupervised learning, inter-view interactions, and association studies.
Throughout this review, we explore various algorithms capable of handling multimodal data, presenting insights into their roles in constructing predictive models for disease detection and classification. The subsequent subsections delve into the presentation of these distinct data types.

4.2 Molecular Data

In this study, we delve into omics data, encompassing extensive datasets acquired through high-throughput techniques. Specifically, we focus on genomics, epigenomics, transcriptomics, and proteomics, as these constitute the omics data types under scrutiny in the reviewed papers.
The majority of the examined studies predominantly leveraged molecular data from two primary databases for disease classification training: the NCI's Genomic Data Commons (GDC) [14] and the Cancer Genome Atlas Program (TCGA) [15–17]. Additionally, the MOSCATO dataset [18] was employed. Primarily utilized for training, this dataset provides a substantial volume of data from a large sample size, offering detailed information on tumor type, grade, RHM score, therapy class, and genomic composition. These characteristics render it particularly suitable for training machine learning models [17].
The Clinical Proteomic Tumor Analysis Consortium (CPTAC) [19], on the other hand, furnishes a public catalog of correlated genomic and proteomic data across various cancer sites, with overlap with TCGA data. The combination of these datasets aims to more comprehensively characterize the molecular state of patients' disease [20].
To introduce epigenomic data [21, 22], we utilize the epigenome-wide association studies (EWAS) dataset [23]. This dataset examines a genome-wide array of quantifiable

epigenetic marks, such as DNA methylation, in different individuals to establish associations between epigenetic variation and specific identifiable phenotypes.
For the integration of transcriptomics data into the model, the studies utilized data obtained through RNA-Seq [24] or RNA-microarrays [25], incorporating mRNA levels for predictive models. In both epigenomics and transcriptomics, the studies commonly relied on the ENCODE database [26].

4.3 Imaging Data

Biomedical research increasingly relies on the integration of molecular and imaging data for a holistic understanding of health and disease. Studies vary in their approaches, with some exclusively leveraging the TCGA database, while others opt for a broader integration of molecular and additional imaging databases.
In the TCGA-focused approach, researchers benefit from its dual repository housing molecular and imaging data. The TCGA database provides a diverse range of imaging data, including CT, MRI, PET, and PET/CT scans, covering both healthy controls and various cancer types.
Alternatively, some studies adopt a more inclusive strategy, integrating molecular databases with supplementary imaging data. Notable sources include private clinical Cancer institutes and the Cancer Imaging Archive (TCIA) [16, 27, 28]. TCIA's repository encompasses images acquired through various techniques, capturing diverse scenarios from healthy controls to different cancer types. The UK BioBank [29, 30] emerges as a valuable resource, contributing a wealth of information, including body and cardiac imaging, genetics, lifestyle measures, biological phenotyping, and health records. The ongoing expansion aims to monitor the health status of up to 100,000 participants over time [16]. In the realm of neuroimaging, two prominent datasets, the Alzheimer's Disease Neuroimaging Initiative (ADNI) [31] and the Enhancing NeuroImaging Genetics through Meta-Analysis (ENIGMA) [32], take the spotlight. Both datasets provide rich repositories of biomarkers, genetics, MRI, and PET imaging data from meticulously tracked patients. However, it's crucial to note that ADNI often limits open access to its data. Consequently, ENIGMA was established to enforce the usage of common, harmonized data analysis, and meta-analysis protocols, fostering free access for collaborators to process their data independently [16].

Partin et al. [33] employed a distinctive methodology, training their model on three key feature types: drug descriptors, gene expression, and histology image tiles. The data collection process involved three primary methodologies.
Firstly, the NCI PDMR [34] played a crucial role by providing comprehensive histopathology assessments, whole-exome sequencing, and RNA-Seq analyses on a subset of tumors. This initial step allowed for the establishment of baseline histology and omic characterization.
In the second phase, transcriptomic data from Patient-Derived Xenografts (PDX), generated through RNA-Seq, underwent transformation into Transcripts Per Kilobase million (TPM). The data were further log2-transformed and standardized for each gene, ensuring a zero mean and a unit standard deviation. Landmark genes, representing significant transcriptomic changes, were meticulously selected from the Library of Integrated Network-Based Cellular Signatures (LINCS) project [35].

Lastly, data were collected from PDX implanted in mice, utilizing 96 PDX models from 89 unique patients. Tumors were subcutaneously grown in NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ (NSG) host mice, staged to an approximate tumor weight of 200 mm^3 for drug studies. Histopathology slides (H&E stained) were digitized at 20x magnification using an Aperio AT2 digital whole slide scanner (Leica Biosystems). A board-certified pathologist ensured consistency with the original patient diagnosis. Regions of interest (ROIs) were annotated using QuPath [36]. Whole slide images underwent processing into individual tiles through the Slideflow software package [37]. Image tiles were extracted from annotated ROIs in a grid pattern at 302 µm by 302 µm with no overlap. Subsequently, they were downsampled to 299 pixels by 299 pixels (10x effective optical magnification). Background tiles were filtered using grayscale, removing those with more than 60% of pixels having a hue value of less than 0.05. To ensure uniformity, digital stain normalization was applied using the Reinhard method [38], followed by standardization to achieve a mean of zero and a variance of one.

5 Methods

The realm of bioinformatics strategies encompasses a diverse array of statistical and machine learning approaches tailored to address the intricacies of omics imaging investigations. This discussion meticulously categorizes general classes of methods, providing a comprehensive overview of their applications and significance. Antonelli et al.'s seminal survey [16] serves as an indispensable foundational resource in this rapidly advancing domain.

In the sphere of supervised classification, a multitude of algorithms has been refined over the years, specifically tailored for the challenges inherent in biomedical contexts. Biomedical data, rich with errors and natural variability in samples, necessitates meticulous algorithm design with tunable parameters. Supervised classification scenarios, where functions assigning labels to data are real-valued, highlight the relevance of regression analysis. This method introduces a regularization term, controlling the count of independent variables. An exemplary technique, the Lasso method, introduces an L1-term to the objective function [39]. Addressing instances with highly correlated features, Canonical Correlation Analysis (CCA) identifies co-expressed features across modalities. Yan et al. [40] further enhance CCA by introducing a penalty term that factors in disease status information, incorporating Laplacian matrices for patients with the same or different diagnoses. The landscape of unsupervised clustering assumes pivotal significance in scenarios involving unlabeled data. Traditional classification or regression algorithms prove impractical, prompting the grouping of similar samples and the segregation of dissimilar ones through clustering methodologies. The lmQCM clustering algorithm, with overlapping groups enabling genes to belong to different clusters, offers a nuanced perspective. Concurrently, hierarchical clustering, as embraced by Diehn et al. [41], unveils common functional or biological themes among genes and identifies modular structures with topological overlaps in networks.

The instrumental role of network analysis in deciphering biological data stems from the intricate connections among cell constituents within an organism. Gene expression

data form the bedrock for constructing networks of interacting genes, with interactions inferred from co-expression. This dynamic landscape extends to bioimaging, where deep learning algorithms emerge as high-performance tools. Renowned for their accuracy in image-related tasks, these algorithms demonstrate success in regulatory genomics, protein classification, protein structure identification, and multi-omics data integration. Recent applications, exemplified by a groundbreaking study mapping tumor gene expression profiles with tumor morphology in pre-operative MRI images, showcase the integration of imaging and omics data. Despite the growing prominence of deep learning, comprehensive tools for integrated analyses of omics and imaging data remain in their nascent stages. Nonetheless, bioinformatics can leverage existing resources for statistical and machine learning analysis on omics data.

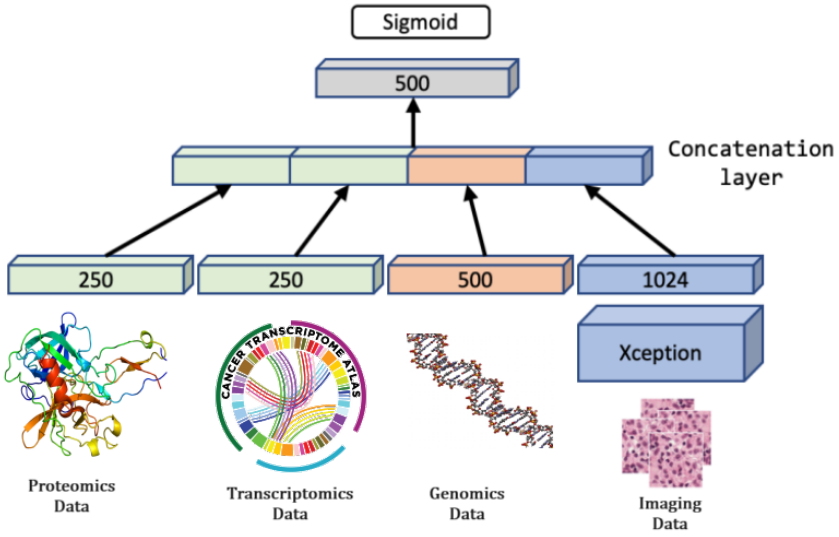

Fig. 3 Example of a Multimodal Network (MM-Net). Adapted from Partin et al.[33]

5.1 Model Description

In investigating the synergy between multimodal AI and the integration of imaging and omics data, a subset of studies has deviated from conventional techniques. A comprehensive review by Shneider et al. [42] has spotlighted endeavors adeptly combining image and omics data through Convolutional Neural Networks (CNNs). These studies notably aimed at enhancing cancer-specific survival (CSS), classifying cancer subtypes, and predicting microsatellite instability (MSI) status. What distinguishes these investigations is their departure from the traditional reliance on either omics or imaging data alone.

One noteworthy study, as highlighted by Shneider et al., delves into the innovative

integration of image and omics data, bringing forth a significant advancement in the field. Chen et al. [43] conducted a comprehensive investigation, leveraging the power of a VGG19-CNN to extract pivotal features from regions of interest within Whole Slide Imaging (WSI) data. Simultaneously, the utilization of self-normalization networks in processing omics data served a dual purpose, not only enhancing feature extraction but also mitigating the risk of overfitting. The true innovation lay in the strategic fusion of these disparate datasets, giving rise to a sophisticated multidimensional array that seamlessly integrated information from both image and molecular domains. This strategic integration marks a paradigm shift, offering a holistic perspective that goes beyond traditional unimodal approaches. Another pioneering effort, as elucidated in the comprehensive review by Shneider et al., underscores the groundbreaking work of Hao et al. [44], who introduced the PAGE-NET model. This model represents a remarkable fusion of genomic data, patient age, and H&E images, showcasing a nuanced approach to data integration. The integration process itself is executed through a patch-wise texture-based CNN, deploying a patch aggregation strategy that captures the intricacies of diverse data modalities. The predictive prowess of the PAGE-NET model is not only limited to its ability to forecast Cancer-Specific Survival (CSS) but extends to its application in a Cox proportional-hazards model. This dual capability not only highlights its versatility but also positions it as a powerful tool in predicting patient outcomes based on a comprehensive understanding of integrated data. The PAGE-NET model, through its innovative design and successful applications, stands as a testament to the potential unlocked by multimodal integration in advancing predictive modeling and clinical insights.
Furthermore, Partin et al. [33] conducted a study using a multimodal neural network (MM-Net) to predict drug response in PDXs. In their comprehensive analysis, MM-Net evaluated six distinct models, each characterized by variations in feature sets and sample types. The unimodal models included UME-Net for gene expression, UMH-Net for histological features, and LGBM (GE) for gene expression. MM-Net, representing a sophisticated multimodal integration approach, seamlessly combined drug descriptors, gene expression data (GE), and histological features. The distinct unimodal models each employed specialized techniques: UME-Net utilized a single-layer neural network for gene expression, UMH-Net employed a Convolutional Neural Network (CNN) for processing histology tiles, and LGBM applied gradient boosting specifically on gene expression data. This fusion into MM-Net showcased not only a unified integration of diverse data modalities but also demonstrated superior performance, particularly in metrics such as the Matthews correlation coefficient (MCC). This outcome underscores the efficacy of multimodal integration as a powerful strategy for advancing the accuracy and reliability of drug response predictions.

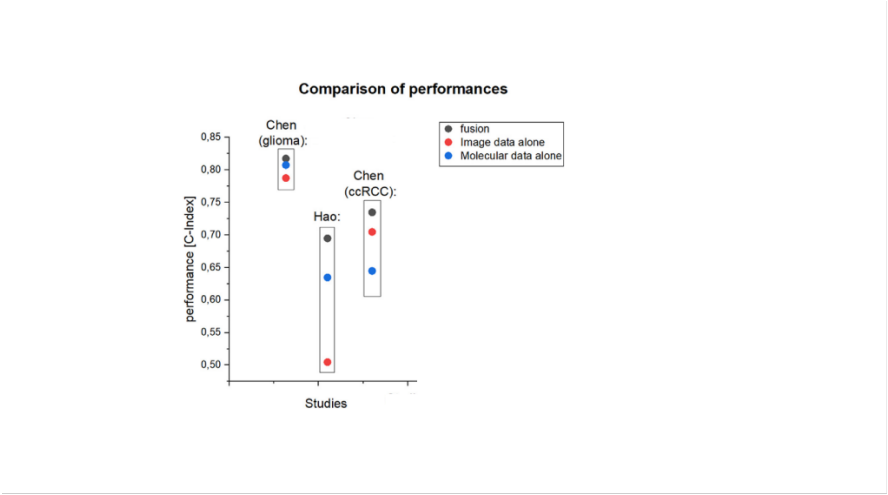

Fig. 4 Comparison of different performance in studies based on C-Index. Adapted from Shneider et al. [42].

6 Discussion

The integration of multimodal AI, harmonizing imaging and omics data, emerges as a transformative force in disease classification and treatment prediction across diverse medical landscapes. Rooted in machine learning and social signal processing, this approach navigates the complex web of biological processes, health indicators, and risk factors, providing a versatile tool in healthcare ([45, 46]).

In the expansive field of cancer research, multimodal AI serves as a catalyst for change, enhancing prognostic accuracy and tailoring treatment strategies. Integrative models, synthesizing insights from omics and histopathological image data, outperform individual factors in various cancers, including liver, lung, renal, and breast cancer [46, 47]. Diseases like head and neck squamous cell carcinoma and neuroendocrine tumors, with their intricate variables, find multimodal AI to be an ideal ally[46, 47]. The potential of AI in predicting treatment outcomes, exemplified in the response to rivastigmine treatment in Alzheimer's disease, underscores its role in personalized medicine and treatment planning [48]. Additionally, AI's rapid diagnosis capabilities, as demonstrated in diseases like necrotizing enterocolitis, highlight its robustness in multivariate analysis [49]. Moving forward in healthcare applications, the development of AI models, including ELMO and MORONET, propels the integration of large-scale AI into biomedical research and healthcare, broadening the scope of multimodal AI applications[50, 51]. Furthermore, AI's utility in predicting and analyzing attribute reduction algorithms in Alzheimer's disease hints at its potential in future clinical and research applications [52]. In the context of cancer, the impact of multimodal AI is profound. Integrative models based on omics and histopathological image data in liver cancer, lung cancer, renal cancer, and breast cancer exhibit superior prognostic accuracy compared to individual factors [46]. The proposed multi-omics data integration model, utilizing UMAP embedding and convolutional neural networks, shows promise

in building a multi-omics prediction model for various cancer types, emphasizing the potential of AI in predicting treatment responses and survival ([53]). This paradigm shift has significant implications for personalized treatment strategies and improved patient outcomes. Transitioning to neurological diseases, multimodal AI offers potential revolutionization in disease prediction and treatment recommendation. Utilizing machine learning for disease prediction and treatment recommendation holds considerable promise for disease prevention, treatment, and management ([54]). Additionally, the development of disease prediction models employing machine learning has the potential to enhance early diagnosis and intervention in neurological diseases such as Alzheimer's. Moreover, the potential of multimodal AI in predicting treatment response and prognosis is exemplified in the context of inflammatory bowel disease (IBD). The identification of predictive biomarkers for therapeutic response highlights the potential of AI-driven approaches to personalize treatment strategies and improve patient outcomes in the management of chronic immune-mediated inflammatory conditions [55].

In conclusion, the integration of multimodal AI, combining imaging and omics data, holds great promise for revolutionizing disease classification and treatment prediction. The comprehensive analysis and predictive capabilities offered by multimodal AI models present opportunities for personalized medicine, improved prognostic accuracy, and enhanced clinical decision-making. As research in this field advances, the transformative potential of multimodal AI in reshaping disease management and patient care becomes increasingly evident, representing the pinnacle of integrating imaging and omics data for improved healthcare outcomes.

References

[1] Bhattacharya, T., Brettin, T., Doroshow, J., Evrard, Y., Greenspan, E., Gryshuk, A., T, H., Lauzon, C., Nissley, D., Penberthy, L., Stahlberg, E., Stevens, R., Streitz, F., Tourassi, G., Xia, F., Zaki, G.: Ai meets exascale computing: advancing cancer research with large-scale high performance computing. Frontiers in Oncology **9** (2019) https://doi.org/10.3389/fonc.2019.00984

[2] Mohsen, F., Ali, H., Hajj, N., Shah, Z.: Artificial intelligence-based methods for fusion of electronic health records and imaging data. Scientific Reports **12** (2022) https://doi.org/10.1038/s41598-022-22514-4

[3] Watson, E., Fard, A., Mar, J.: Computational methods for single-cell imaging and omics data integration. Frontiers in Molecular Biosciences **8** (2022) https://doi.org/10.3389/fmolb.2021.768106

[4] Karim, R., Islam, T., Lange, C., Rebholz-Schuhmann, D., Decker, S.: Adversary-aware multimodal neural networks for cancer susceptibility prediction from multiomics data. Ieee Access **10**, 54386–54409 (2022) https://doi.org/10.1109/access.2022.3175816

[5] Chaudhary, K., Poirion, O., Lu, L.: Deep learning–based multi-omics integration robustly predicts survival in liver cancer. Clinical Cancer Research **24**, 1248–1259 (2018) https://doi.org/10.1158/1078-0432.ccr-17-0853

[6] Sone, K., Toyohara, Y., Taguchi, A., Miyamoto, Y., Uchino-Mori, M., Iriyama, T., Tsuruga, T., Osuga, Y.: Application of artificial intelligence in gynecologic malignancies: a review. Journal of Obstetrics and Gynaecology Research **47**, 2577–2585 (2021) https://doi.org/10.1111/jog.14818

[7] Ramazzotti, D., A, L., Wang, B., Batzoglou, S., Sidow, A.: Multi-omic tumor data reveal diversity of molecular mechanisms that correlate with survival. Nature Communications **9** (2018) https://doi.org/10.1038/s41467-018-06921-8

[8] Arimura, H., Soufi, M., Kamezawa, H., Ninomiya, K., Yamada, M.: Radiomics with artificial intelligence for precision medicine in radiation therapy. Journal of Radiation Research **60**, 150–157 (2018) https://doi.org/10.1093/jrr/rry077

[9] Efremova, M., Teichmann, S.: Computational methods for single-cell omics across modalities. Nature Chemical Biology **17**, 14–17 (2020) https://doi.org/10.1038/s41592-019-0692-4

[10] PET/CT Scan | University Radiology Associates, LLP | SUNY Upstate Medical University. https://www.upstate.edu/ura/pet-ct-scan.php Accessed 2023-12-27

[11] H&E stain. Page Version ID: 1191484566 (2023). https://en.wikipedia.org/w/index.php?title=H%26E_stain&oldid=1191484566 Accessed 2023-12-27

[12] Using FAST on Whole-Slide Images (WSI) | FAST | Documentation. https://fast.eriksmistad.no/python-tutorial-wsi.html Accessed 2023-12-27

[13] Supporting Multi-omics Approaches - IT. https://www.thermofisher.com/uk/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/spotlight-articles/supporting-multi-omics-approaches.html Accessed 2023-12-22

[14] Home | NCI Genomic Data Commons. https://gdc.cancer.gov/ Accessed 2023-11-30

[15] The Cancer Genome Atlas Program (TCGA) - NCI. Archive Location: nciglobal,ncienterprise (2022). https://www.cancer.gov/ccg/research/genome-sequencing/tcga Accessed 2023-11-29

[16] Antonelli, L., Guarracino, M.R., Maddalena, L., Sangiovanni, M.: Integrating imaging and omics data: A review. Biomedical Signal Processing and Control **52**, 264–280 (2019) https://doi.org/10.1016/j.bspc.2019.04.032 . Accessed 2023-11-29

[17] Sun, R., Limkin, E.J., Vakalopoulou, M., Dercle, L., Champiat, S., Han, S.R., Verlingue, L., Brandao, D., Lancia, A., Ammari, S., Hollebecque, A., Scoazec, J.-Y., Marabelle, A., Massard, C., Soria, J.-C., Robert, C., Paragios, N., Deutsch, E., Ferté, C.: A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. The Lancet Oncology **19**(9), 1180–1191 (2018) https://doi.org/10.1016/S1470-2045(18)30413-3 . Accessed 2023-11-29

[18] Massard, C., Michiels, S., Ferté, C., Le Deley, M.-C., Lacroix, L., Hollebecque, A., Verlingue, L., Ileana, E., Rosellini, S., Ammari, S., Ngo-Camus, M., Bahleda, R., Gazzah, A., Varga, A., Postel-Vinay, S., Loriot, Y., Even, C., Breuskin, I., Auger, N., Job, B., De Baere, T., Deschamps, F., Vielh, P., Scoazec, J.-Y., Lazar, V., Richon, C., Ribrag, V., Deutsch, E., Angevin, E., Vassal, G., Eggermont, A., André, F., Soria, J.-C.: High-Throughput Genomics and Clinical Outcome in Hard-to-Treat Advanced Cancers: Results of the MOSCATO 01 Trial. Cancer Discovery **7**(6), 586–595 (2017) https://doi.org/10.1158/2159-8290.CD-16-1396 . Accessed 2023-11-28

[19] CPTAC | Office of Cancer Clinical Proteomics Research. https://proteomics.cancer.gov/programs/cptac Accessed 2023-12-01

[20] Boehm, K.M., Khosravi, P., Vanguri, R., Gao, J., Shah, S.P.: Harnessing multimodal data integration to advance precision oncology. Nature Reviews Cancer **22**(2), 114–126 (2022) https://doi.org/10.1038/s41568-021-00408-3 . Number: 2 Publisher: Nature Publishing Group. Accessed 2023-11-29

[21] Tabakhi, S., Suvon, M.N.I., Ahadian, P., Lu, H.: Multimodal Learning for Multi-omics: A Survey. World Scientific Annual Review of Artificial Intelligence **01**,

2250004 (2023) https://doi.org/10.1142/S2811032322500047 . Accessed 2023-11-29

[22] Sun, D., Li, A., Tang, B., Wang, M.: Integrating genomic data and pathological images to effectively predict breast cancer clinical outcome. Computer Methods and Programs in Biomedicine **161**, 45–53 (2018) https://doi.org/10.1016/j.cmpb.2018.04.008 . Accessed 2023-12-01

[23] EWAS Datahub. https://ngdc.cncb.ac.cn/ewas/datahub Accessed 2023-12-01

[24] Wang, Z., Gerstein, M.B., Snyder, M.: Rna-seq: a revolutionary tool for transcriptomics. Nature Reviews Genetics **10**, 57–63 (2009)

[25] Sealfon, S.C., Chu, T.T.: Rna and dna microarrays. Methods in molecular biology **671**, 3–34 (2011)

[26] ENCODE. https://www.encodeproject.org/ Accessed 2023-12-01

[27] The Cancer Imaging Archive (TCIA). https://imaging.cancer.gov/informatics/cancer_imaging_archive.htm Accessed 2023-11-30

[28] Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L.H., Aerts, H.J.W.L.: Artificial intelligence in radiology. Nature reviews. Cancer **18**(8), 500–510 (2018) https://doi.org/10.1038/s41568-018-0016-5 . Accessed 2023-11-28

[29] UK Biobank - UK Biobank. https://www.ukbiobank.ac.uk/ Accessed 2023-12-06

[30] Gibson, L.M., Nolan, J., Littlejohns, T.J., Mathieu, E., Garratt, S., Doherty, N., Petersen, S., Harvey, N.C.W., Sellors, J., Allen, N.E., Wardlaw, J.M., Jackson, C.A., Sudlow, C.L.M.: Factors associated with potentially serious incidental findings and with serious final diagnoses on multi-modal imaging in the UK Biobank Imaging Study: A prospective cohort study. PLOS ONE **14**(6), 0218267 (2019) https://doi.org/10.1371/journal.pone.0218267 . Accessed 2023-12-06

[31] ADNI | Alzheimer's Disease Neuroimaging Initiative. https://adni.loni.usc.edu/ Accessed 2023-12-06

[32] ENIGMA. https://enigma.ini.usc.edu/ Accessed 2023-12-06

[33] Partin, A., Brettin, T., Zhu, Y., Dolezal, J.M., Kochanny, S., Pearson, A.T., Shukla, M., Evrard, Y.A., Doroshow, J.H., Stevens, R.L.: Data augmentation and multimodal learning for predicting drug response in patient-derived xenografts from gene expressions and histology images. arXiv. arXiv:2204.11678 [q-bio] (2022). http://arxiv.org/abs/2204.11678 Accessed 2023-11-28

[34] Patient-Derived Models Repository (PDMR). https://pdmr.cancer.gov/ Accessed 2023-11-28

[35] NIH LINCS Program. https://lincsproject.org/ Accessed 2023-11-28

[36] Welcome to QuPath! — QuPath 0.4.4 documentation. https://qupath.readthedocs.io/en/stable/ Accessed 2023-11-30

[37] Dolezal, J.M., Kochanny, S., Dyer, E., Srisuwananukorn, A., Sacco, M., Howard, F.M., Li, A., Mohan, P., Pearson, A.T.: Slideflow: Deep Learning for Digital Histopathology with Real-Time Whole-Slide Visualization. arXiv. arXiv:2304.04142 [cs, eess, q-bio] (2023). https://doi.org/10.48550/arXiv.2304.04142 . http://arxiv.org/abs/2304.04142 Accessed 2023-11-30

[38] Reinhard, E., Ashikhmin, M., Gooch, B., Shirley, P.: Color Transfer between Images. IEEE Computer Graphics and Applications **21**, 34–41 (2001) https://doi.org/10.1109/38.946629

[39] Tibshirani, R.: Regression Shrinkage and Selection Via the Lasso. Journal of the Royal Statistical Society: Series B (Methodological) **58**(1), 267–288 (1996) https://doi.org/10.1111/j.2517-6161.1996.tb02080.x . _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.2517-6161.1996.tb02080.x. Accessed 2023-12-12

[40] Yan, J., Risacher, S.L., Nho, K., Saykin, A.J., Shen, L.: IDENTIFICATION OF DISCRIMINATIVE IMAGING PROTEOMICS ASSOCIATIONS IN ALZHEIMER'S DISEASE VIA A NOVEL SPARSE CORRELATION MODEL. In: Biocomputing 2017, pp. 94–104. WORLD SCIENTIFIC, ??? (2016). https://doi.org/10.1142/9789813207813_0010 . https://www.worldscientific.com/doi/abs/10.1142/9789813207813_0010 Accessed 2023-12-12

[41] Airoldi, E.M., Erosheva, E.A., Fienberg, S.E., Joutard, C., Love, T., Shringarpure, S.: Reconceptualizing the classification of PNAS articles. Proceedings of the National Academy of Sciences of the United States of America **107**(49), 20899–20904 (2010) https://doi.org/10.1073/pnas.1013452107 . Accessed 2023-12-12

[42] Integration of deep learning-based image analysis and genomic data in cancer pathology: A systematic review. European Journal of Cancer **160**, 80–91 (2022) https://doi.org/10.1016/j.ejca.2021.10.007 . Publisher: Pergamon. Accessed 2023-11-28

[43] Chen, R.J., Lu, M.Y., Wang, J., Williamson, D.F.K., Rodig, S.J., Lindeman, N.I., Mahmood, F.: Pathomic fusion: An integrated framework for fusing histopathology and genomic features for cancer diagnosis and prognosis. IEEE Transactions on Medical Imaging **41**(4), 757–770 (2022) https://doi.org/10.1109/TMI.2020.3021387

[44] Hao, J., Kosaraju, S.C., Tsaku, N.Z., Song, D.H., Kang, M.: PAGE-Net: Interpretable and integrative deep learning for survival analysis using histopathological images and genomic data. Pacific Symposium on Biocomputing **25**(2020),

355–366 (2020)

[45] Cukurova, M., Kent, C., Luckin, R.: Artificial intelligence and multimodal data in the service of human decision-making: a case study in debate tutoring. British Journal of Educational Technology **50**, 3032–3046 (2019) https://doi.org/10.1111/bjet.12829

[46] Zeng, H., Chen, L., Huang, Y., Luo, Y., Ma, X.: Integrative models of histopathological image features and omics data predict survival in head and neck squamous cell carcinoma. Frontiers in Cell and Developmental Biology **8** (2020) https://doi.org/10.3389/fcell.2020.553099

[47] Puliani, G., Vito, V.D., Feola, T., Sesti, F., Centello, R., Pandozzi, C., Tarsitano, M.G., Verrico, M., Lenzi, A., Isidori, A.M., Giannetta, E., Faggiano, A.: Netest: a systematic review focusing on the prognostic and predictive role. Neuroendocrinology **112**, 523–536 (2021) https://doi.org/10.1159/000518873

[48] Farlow, M.R., Hake, A.M., Messina, J., Hartman, R.D., Veach, J., Anand, R.: Response of patients with alzheimer disease to rivastigmine treatment is predicted by the rate of disease progression. Archives of Neurology **58** (2001) https://doi.org/10.1001/archneur.58.3.417

[49] Gao, W., Pei, Y., Liang, H., Lv, J., Chen, J., Zhong, W.: Multimodal ai system for the rapid diagnosis and surgical prediction of necrotizing enterocolitis. IEEE Access **9**, 51050–51064 (2021) https://doi.org/10.1109/access.2021.3069191

[50] Zhang, Y., Shi, R., Chen, C., Duan, M., Liu, S., Ren, Y., Huang, L., Dai, X., Zhou, F.: Elmo: an efficient logistic regression-based multi-omic integrated analysis method for breast cancer intrinsic subtypes. IEEE Access **8**, 5121–5130 (2020) https://doi.org/10.1109/access.2019.2960373

[51] Shao, W., Huang, Z., Tang, H., Zhang, J., Ding, Z., Huang, K.: Moronet: multi-omics integration via graph convolutional networks for biomedical data classification (2020) https://doi.org/10.1101/2020.07.02.184705

[52] Li, W.: Prediction and analysis of attribute reduction algorithm in rough set in alzheimer's disease (2023) https://doi.org/10.1117/12.2680727

[53] Elkarami, B., Alkhateeb, A., Qattous, H., Alshomali, L., Shahrrava, B.: Multi-omics data integration model based on umap embedding and convolutional neural network. Cancer Informatics **21**, 117693512211242 (2022) https://doi.org/10.1177/11769351221124205

[54] Mahata, S., Kapadiya, Y.B., Kushwaha, V., Joshi, V., Farooqui, Y.: Disease prediction and treatment recommendation using machine learning. International Journal for Research in Applied Science and Engineering Technology **11**, 1232–1237 (2023) https://doi.org/10.22214/ijraset.2023.49641

[55] Elhag, D.A., Kumar, M., Saadaoui, M., Akobeng, A.K., Al-Mudahka, F., Elawad, M., Al Khodor, S.: Inflammatory Bowel Disease Treatments and Predictive Biomarkers of Therapeutic Response. International Journal of Molecular Sciences **23**(13), 6966 (2022) https://doi.org/10.3390/ijms23136966 . Accessed 2023-12-22