HHS Public Access

Author manuscript

Adv Cancer Res. Author manuscript; available in PMC 2024 September 27.

Published in final edited form as:

Adv Cancer Res. 2024 ; 163: 1–38. doi:10.1016/bs.acr.2024.08.001.

Deep learning-based multimodal spatial transcriptomics analysis for cancer

Pankaj Rajdeo[a], **Bruce Aronow**[a,b], **V.B. Surya Prasath**[a,b,c,d,*]

[a]Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States

[b]Department of Pediatrics, College of Medicine, University of Cincinnati, Cincinnati, OH, United States

[c]Department of Biomedical Informatics, College of Medicine, University of Cincinnati, Cincinnati, OH, United States

[d]Department of Computer Science, University of Cincinnati, Cincinnati, OH, United States

Abstract

The advent of deep learning (DL) and multimodal spatial transcriptomics (ST) has revolutionized cancer research, offering unprecedented insights into tumor biology. This book chapter explores the integration of DL with ST to advance cancer diagnostics, treatment planning, and precision medicine. DL, a subset of artificial intelligence, employs neural networks to model complex patterns in vast datasets, significantly enhancing diagnostic and treatment applications. In oncology, convolutional neural networks excel in image classification, segmentation, and tumor volume analysis, essential for identifying tumors and optimizing radiotherapy.

The chapter also delves into multimodal data analysis, which integrates genomic, proteomic, imaging, and clinical data to offer a holistic understanding of cancer biology. Leveraging diverse data sources, researchers can uncover intricate details of tumor heterogeneity, microenvironment interactions, and treatment responses. Examples include integrating MRI data with genomic profiles for accurate glioma grading and combining proteomic and clinical data to uncover drug resistance mechanisms.

DL's integration with multimodal data enables comprehensive and actionable insights for cancer diagnosis and treatment. The synergy between DL models and multimodal data analysis enhances diagnostic accuracy, personalized treatment planning, and prognostic modeling. Notable applications include ST, which maps gene expression patterns within tissue contexts, providing critical insights into tumor heterogeneity and potential therapeutic targets.

In summary, the integration of DL and multimodal ST represents a paradigm shift towards more precise and personalized oncology. This chapter elucidates the methodologies and applications of these advanced technologies, highlighting their transformative potential in cancer research and clinical practice.

*Corresponding author. surya.prasath@cchmc.org.

1. Introduction

1.1 Deep learning and its significance in oncology

Deep learning (DL), a subset of machine learning, uses artificial neural networks to model complex patterns in data. In oncology, this technology has revolutionized the analysis of vast datasets, including medical images, genomic data, and clinical records, enhancing applications from diagnosis to treatment planning. These capabilities significantly improve cancer diagnosis, treatment planning, and prognostication (Li, Jiang, Zhang, & Zhu, 2023; Lipkova et al., 2022; Steyaert et al., 2023b).

For instance, convolutional neural networks (CNNs) excel in tasks such as image classification, segmentation, and tumor volume segmentation. These capabilities are crucial for accurately identifying tumors and optimizing radiotherapy planning. By enabling precise mapping of tumors, CNNs facilitate a shift towards precision medicine, where treatment strategies are meticulously tailored to the individual characteristics of each patient's tumor, thereby significantly enhancing treatment outcomes (Sharma, Nayak, Balabantaray, Tanveer, & Nayak, 2024).

Additionally, advancements in neural network architectures, particularly through the use of transfer learning and data augmentation, have tailored DL tools to specific oncology needs, enhancing both research and clinical applications.

To illustrate these advancements, Table 1 summarizes some studies that demonstrate DL's transformative role in oncology:

DL's continued integration has significantly enhanced the accuracy and efficiency of cancer care. This is particularly evident in the use of CNNs for precise tumor segmentation and the application of transfer learning. Transfer learning involves adapting pre-trained models —originally developed for one task—to new, related tasks. This approach is highly effective in oncology, where models trained on large, diverse datasets are fine-tuned to recognize specific cancer types. By using transfer learning, clinicians can leverage the knowledge gained from vast amounts of existing data, thereby reducing the need for extensive labeled datasets specific to each cancer type. This accelerates the development of diagnostic tools and enables more rapid and accurate customization of treatment plans, which helps further the goals of the precision medicine (Aneja, Aneja, Abas, & Naim, 2021; Hanczar, Bourgeais, & Zehraoui, 2022; Luo & Bocklitz, 2023).

Ultimately, the significance of multimodal data analysis cannot be overstated. It represents a substantial shift towards more comprehensive and integrated methods in cancer research and treatment, which is the primary focus of this chapter.

1.2 Multimodal data analysis in cancer research

Multimodal data analysis involves integrating diverse data types—such as genomic, proteomic, imaging, and clinical data—to gain a holistic understanding of complex biological systems (Acosta, Falcone, Rajpurkar, & Topol, 2022). This foundational approach

enhances the depth and accuracy of insights into cancer biology by leveraging the strengths of varied data sources.

- **Genomic data** provide insights into genetic mutations and variations.
- **Proteomic data** reveal details about protein expression and modifications.
- **Imaging data** offer a detailed view of tumor morphology and microenvironment.
- **Clinical data** provide context about patient history and treatment outcomes.

For a more detailed look at the practical applications of these integrations in cancer research, particularly in the context of spatial transcriptomics (ST), see Section 2.2.

1.2.1 Key examples and benefits of integrating various data types

- **Genomic and imaging data integration:** This integration enhances tumor subtype identification and treatment response prediction. For example, integrating MRI data with genomic profiles can distinguish between different glioma grades, leading to more accurate treatment decisions (Cebula et al., 2020).
- **Proteomic and clinical data integration:** This combination provides insights into drug resistance mechanisms and potential therapeutic targets. It also helps monitor disease progression and evaluate treatment effectiveness in real-time (Price et al., 2022).
- **Genomic, proteomic, and imaging data integration:** This comprehensive approach uncovers complex interactions between genes, proteins, and the tumor microenvironment (TME), facilitating the development of multi-targeted therapies and improving patient outcomes (Boehm, Khosravi, Vanguri, Gao, & Shah, 2022).

Refer to Table 2 for additional examples of multimodal data integration in cancer research.

1.3 Synergistic potential of deep learning and multimodal approaches

The synergy between DL and multimodal data analysis holds immense potential for advancing precision medicine and enhancing patient outcomes in oncology. CNNs, in particular, excel in processing and interpreting complex medical data, such as imaging, genomics, and clinical records.

One notable application is the use of ST, which allows the mapping of gene expression patterns within the spatial context of tissues. This provides critical insights into tumor heterogeneity, microenvironment interactions, and treatment response. Combining this with multimodal approaches, which incorporate various data types like genomic, proteomic, and imaging data, enables DL models to uncover more comprehensive and actionable information for cancer diagnosis and treatment (Halawani, Buchert, & Chen, 2023; Hu, Sajid, Lv, Liu, & Sun, 2022; Li, Li, You, Wei, & Xu, 2023).

Fig. 1 illustrates the comprehensive workflow for analyzing cancer using DL-based multimodal ST:

This figure illustrates the comprehensive workflow for analyzing cancer using multimodal ST and DL.

- **A.** Tissue sample preparation, including sectioning, staining, and ST profiling.
- **B.** Data acquisition phase, capturing ST data, imaging data (e.g., hematoxylin and eosin [H&E] staining), and optional genomic/proteomic data.
- **C.** Preprocessing and integration of multimodal data, ensuring consistency and compatibility.
- **D.** Application of DL models, such as CNNs, autoencoders, transformers, and GANs, for cell type identification, spatial domain prediction, gene expression pattern analysis, and integration of multi-omics features.
- **E.** Biological interpretation of results, revealing insights into tumor heterogeneity, tumor-microenvironment interactions, and potential therapeutic targets, advancing precision oncology.

1.3.1 Benefits of this synergy—

Enhanced diagnostic accuracy is achieved through DL models that analyze vast imaging data volumes to identify subtle cancer patterns, which might be missed by human observers. By integrating genomic and proteomic data, these models correlate molecular signatures with imaging findings, leading to more accurate diagnoses (Jiang, Hu, Wang, & Zhang, 2023). Personalized treatment planning benefits from multimodal data providing a holistic view of a patient's condition, enabling DL models to predict treatment responses and outcomes more effectively. For example, combining imaging data with genomic profiles can tailor chemotherapy or radiotherapy plans to the individual patient's tumor characteristics, increasing treatment efficacy and reducing adverse effects (Joo et al., 2021). Improved prognostic models emerge from DL models that integrate longitudinal data from various sources to predict disease progression and patient survival more accurately. This capability is crucial for developing personalized follow-up strategies and improving long-term patient management (Cascarano et al., 2023). Real-time decision support is facilitated by the synergy between DL and multimodal approaches, allowing real-time clinical decision support during surgery or radiotherapy. Real-time data from imaging modalities processed by DL algorithms guide procedures, ensuring precise targeting of cancerous tissues while sparing healthy ones (Steyaert et al., 2023b). Accelerated research and development are achieved by harnessing DL to analyze multimodal datasets, helping researchers uncover new biomarkers and therapeutic targets more efficiently. This acceleration leads to faster development of innovative treatments and diagnostics, ultimately benefiting patients (Johnson et al., 2021).

The synergy between DL and multimodal data analysis, especially through the integration of ST, signifies a paradigm shift in oncology. This transition is paving the way for a new era of precision medicine characterized by enhanced diagnostic accuracy, more personalized treatment strategies, and overall improved patient outcomes.

Having explored the synergistic potential of DL and multimodal approaches, we now have established a solid foundation for understanding how these advanced methodologies are

revolutionizing the field of oncology. Next, we will delve deeper into the core concepts of multimodal ST. This cutting-edge area further exemplifies the transformative power of integrating diverse data types to unravel the complexities of cancer, which is offering new avenues for research and treatment that promise to enhance the precision and effectiveness of oncology therapies.

2. Core concepts of multimodal spatial transcriptomics

2.1 Basics of spatial transcriptomics

ST is a transformative technique in molecular biology that allows for the visualization and quantification of gene expression within tissue sections, preserving the spatial context of the cells. This innovative approach provides essential insights into the variability of gene expression across different regions of tissues, enhancing our understanding of cellular functions and tissue organization in both health and disease (Ståhl et al., 2016).

Key principles of ST include spatial resolution, which maintains spatial information about where gene expression occurs within the tissue, providing crucial context for understanding cell interactions within their microenvironments and how these interactions influence overall tissue function. High-throughput sequencing is utilized by ST to analyze the transcriptome across thousands of spatially resolved locations within a tissue sample, enabling a comprehensive analysis of gene expression patterns. Additionally, the integration with imaging combines transcriptomic data with high-resolution tissue imaging, creating a detailed map of gene expression patterns corresponding to specific histological features, which helps to correlate morphological characteristics with molecular data, enhancing the precision of biological insights.

2.1.1 Methods

- **In situ hybridization:** Techniques like MERFISH and seqFISH use fluorescent probes to detect RNA molecules within tissues, allowing for high spatial resolution localization of gene expression (Du et al., 2023).
- **Spatially resolved transcriptomics platforms:** Technologies like 10x Genomics' Visium and Slide-seq use spatial barcoding to map gene expression. These platforms facilitate precise mapping of RNA molecules to their spatial coordinates within the tissue (Williams, Lee, Asatsuma, Vento-Tormo, & Haque, 2022).
- **Laser capture microdissection (LCM):** LCM involves using a laser to precisely cut out specific regions of tissue for RNA extraction and sequencing. This method provides high spatial resolution and is particularly useful for analyzing small or rare cell populations within a tissue (Emmert-Buck et al., 1996).

2.1.2 Importance in mapping the transcriptome to tissue locations—

Mapping the transcriptome to specific tissue locations is transformative for several reasons (Williams et al., 2022). Understanding tissue heterogeneity is enhanced as ST allows researchers to identify distinct gene expression profiles across different tissue regions, shedding

light on how various cell types contribute to tissue function and pathology. In terms of disease mechanisms, ST reveals interactions between cancer cells and surrounding stromal and immune cells, providing insights into tumor growth and metastasis mechanisms. In developmental biology, ST is invaluable for studying gene expression during tissue development, aiding in understanding the processes of tissue formation, differentiation, and organization.

Case studies illustrate the application of ST in uncovering complex cellular interactions and identifying therapeutic targets in diseases, particularly in cancer research. In colorectal cancer, a recent study by Peng et al. (2023b) utilized ST to explore the interactions between fibroblasts and myeloid cells. By integrating single-cell RNA sequencing, ST, and bulk RNA sequencing data, the researchers identified a pro-tumorigenic interaction between MFAP5+ fibroblasts and C1QC+ macrophages. These interactions were mapped spatially within the tumor, highlighting specific signaling pathways that contribute to the malignant behavior of colorectal cancer. This study underscores the importance of ST in elucidating the complex cellular interactions within the TME and identifying potential therapeutic targets. In glioblastoma, another example is the work by Liu et al. (2023), who integrated single-cell RNA sequencing with ST to analyze the cellular heterogeneity. They identified distinct clusters of malignant cells with unique transcriptional and functional properties. ST allowed them to map these clusters within the tumor, revealing their spatial colocalization and interactions with the TME. This integrated approach provided novel insights into the mechanisms of tumor progression and resistance to therapy in glioblastoma.

These examples underscore how ST, by preserving the spatial context of gene expression, helps deepen our understanding of tissue architecture and cellular interactions, which is crucial for developing targeted therapies and improving clinical outcomes.

The following case studies illustrate the application of ST in uncovering complex cellular interactions and identifying therapeutic targets in diseases, particularly in cancer research.

2.2 Multimodal spatial transcriptomics: integrating varied data types

As discussed in Section 1.3, multimodal ST marks a significant advancement in studying complex biological systems, especially in the context of cancer research. By integrating diverse data types such as genomics, proteomics, imaging, and clinical data, researchers can achieve a holistic view of disease mechanisms. This section focuses on how these integrations are uniquely beneficial in the context of ST.

2.2.1 Integrating genomics and proteomics with spatial transcriptomics

Genomic data integration: Combining ST with genomic data offers a comprehensive view of the genetic landscape within tissues. This integration is crucial for revealing mutations and structural variations that drive cancer progression. Tools like MethCNA integrate DNA methylation and copy number variation data, aiding in the identification of oncogenic drivers within their spatial context in tissues (Deng, Yang, Zhang, Xiao, & Cai, 2018). For example, integrating ST with genomic perturbation data enhances gene regulatory network inference, allowing researchers to predict how specific genomic alterations impact gene expression

across different tissue regions (Liang, Young, Hung, Raftery, & Yeung, 2019). This helps in understanding how mutations within the same tumor can lead to varied phenotypic outcomes.

Proteomic data integration: Proteomics complements ST by providing insights into protein expression and modifications, the functional executors of genetic information. Proteogenomic approaches integrate patient-specific genomic and proteomic data to characterize the mutational landscape and protein signaling pathways in tumors, identifying patient-specific drug targets and resistance mechanisms (Schmitt et al., 2021).

2.2.2 Complementing ST with imaging and clinical data

Imaging data integration: Integrating imaging data with ST allows for correlating morphological features with molecular changes, enhancing the understanding of tumor phenotypes and treatment responses. For instance, the FDTrans model developed by Cai et al. (2023) integrates histopathological images with genomic data to classify lung cancer subtypes, demonstrating the potential of this approach.

Clinical data integration: Combining ST with clinical data helps identify biomarkers predictive of therapy response and disease prognosis. Integrating clinical information with molecular profiles, as done by Wilson, Li, Yu, Kuan, and Wang (2019) identifies gene sets relevant to prognostic predictions across cancer types.

Fig. 2 illustrates the application of ST techniques in a cancer tissue section, highlighting the detailed visualization methods used to study the TME.

a. **Multiplexed immunofluorescence image:** This image shows the spatial distribution of three different proteins within the tissue, each protein is denoted by a specific color.

b. **H&E stained tissue section:** Displays the overall tissue morphology and cellular structure, providing a visual context for the underlying tissue architecture.

c. **ST heatmap:** This heatmap demonstrates the relative expression levels of a gene of interest across different regions of the tissue, indicating areas of high and low expression.

d. **Integrated view:** This circle contains a zoomed-in view of the H&E-stained tissue, overlaid with color-coded indicators for tumor cells (yellow), immune cells (red), and stromal cells (green and purple), illustrating the cellular composition and interaction within the microenvironment.

2.2.3 Successful multimodal ST studies in cancer research—

The following case studies demonstrate the successful application of multimodal ST in cancer research, providing insights that guide the development of more effective and personalized treatment strategies. Harikumar, Quinn, Rana, Gupta, and Venkatesh (2021) proposed a framework to predict gene responses to drugs using patient-specific single-cell expression data and population-level drug-response data. This dual-channel approach integrates single-cell RNA-seq with drug response data, providing personalized predictions of drug efficacy. This

framework has been applied to glioblastoma, illustrating how multimodal integration can enhance personalized treatment strategies.

Another notable example is the FDTrans model by Cai et al. (2023), which exemplifies successful multimodal integration by combining histopathological images with genomic data to classify lung cancer subtypes. This model achieved high accuracy and AUC scores, demonstrating the potential of integrating imaging and genomic data to improve diagnostic precision and treatment planning. Similarly, Schmitt et al. (2021) used proteogenomics to study the mutational landscape of melanoma patients. By integrating genomic and proteomic data, they identified pathways associated with melanoma development and immunotherapy response, highlighting the importance of combining multiple data types to uncover molecular mechanisms and identify therapeutic targets.

In a comprehensive study by Wei, Zhang, Weng, Chen, and Cai (2021), various computational methods for integrating multi-omics data across pan-cancer datasets were assessed. This study demonstrated the ability of integrative methods to identify distinct tumor compositions and molecular patterns, providing valuable insights for pan-cancer analysis and highlighting the importance of multimodal data integration in cancer research. In summary, multimodal ST provides a comprehensive framework for understanding the complex molecular landscape of cancers, enhancing our knowledge of tumor biology and paving the way for personalized medicine.

2.3 Unraveling cancer complexity with multimodal spatial transcriptomics

Multimodal ST empowers researchers to dissect tumor heterogeneity, analyze the TME, and evaluate therapy responses with unprecedented detail and precision. These capabilities lead to several key insights.

2.3.1 Key insights from multimodal ST

Tumor heterogeneity: One of the primary insights gained from multimodal ST is the detailed characterization of tumor heterogeneity. Tumors are composed of a mosaic of genetically distinct cell populations, each contributing differently to disease progression and therapy resistance. ST, combined with single-cell sequencing, allows for the mapping of these heterogeneous cell populations within their spatial context. For instance, Halawani et al. (2023) highlighted how DL frameworks applied to single-cell and ST data can unravel the complex gene transcription profiles and mutation spectra within tumors. This approach has been crucial in identifying subclonal populations that drive disease progression and metastasis. In esophageal cancer, the integration of multi-omics data has provided insights into the cellular and genetic heterogeneity within tumors. Li et al. (2023) discussed how artificial intelligence (AI) can analyze and interpret multi-omics data, which enables a comprehensive understanding of tumor heterogeneity and its implications for disease progression and treatment strategies. This high-dimensional characterization helps in identifying novel cell types and understanding their roles in tumor biology.

TME: The TME plays a pivotal role in cancer development and therapy response. Multimodal ST allows for the spatial profiling of immune and stromal cells within

the TME, providing insights into their interactions with tumor cells. Hu et al. (2022) emphasized the importance of spatial profiling technologies in evaluating the transcriptional activity of immune and stromal cells within the TME. This multidimensional classification facilitates the study of tumor-immune interactions, the evolutionary trajectory of tumors, and the identification of immune evasion mechanisms. A study by Zhang et al. (2024a) on hepatocellular carcinoma revealed how single-cell sequencing and ST can uncover the spatial distribution and functional states of tumor-infiltrating lymphocytes. The identification of stimulatory dendritic cells and macrophages as potential biomarkers underscores the importance of understanding the spatial organization of immune cells in predicting therapy responses.

Therapy response: Understanding the spatial and molecular determinants of therapy response is crucial for developing effective cancer treatments. Multimodal ST provides a platform for identifying biomarkers and therapeutic targets by linking molecular profiles to spatial contexts. For example, Lyubetskaya et al. (2022) demonstrated how ST can capture key tumor features such as hypoxia, necrosis, and vasculature, which are critical for understanding therapy responses. The ability to map these features within the tumor landscape helps in identifying regions of therapeutic resistance and potential targets for intervention. In breast cancer, Bottosso et al. (2024) highlighted the use of precision medicine to predict drug sensitivity based on a deeper molecular understanding of the disease. The integration of genomic, transcriptomic, and imaging data allows for the identification of biomarkers that predict response to specific therapies and can aid in the personalization of treatment strategies.

Having established the foundational concepts and diverse applications of multimodal ST, we now shift our focus to the innovative role of DL in this domain. This transition marks a pivotal move towards harnessing sophisticated computational models to further unravel the complexities of cancer at a molecular and spatial level.

3. Deep learning approaches in multimodal spatial transcriptomics

Building upon our exploration of multimodal ST, Section 3 delves into the transformative role of DL in refining and advancing this field. As depicted in Fig. 3, these models integrate diverse data types—including genomic, proteomic, imaging, clinical, and ST—to enable comprehensive analysis and interpretation of cancer data. Here, we focus on cutting-edge deep learning models that bring unprecedented precision to analyzing ST data. These models not only interpret complex biological relationships within tissues but also push the boundaries of diagnostic and prognostic capabilities, marking a significant evolution in the tools available to oncologists and researchers.

This figure illustrates the process of multimodal data fusion and the application of DL to achieve improved diagnostic accuracy, personalized treatments, and a better understanding of tumor heterogeneity and the TME.

3.1 Deep learning models for spatial transcriptomics data

ST technologies provide a detailed map of gene expression patterns within the spatial architecture of tissues. The complexity of this data necessitates the use of advanced DL models that can identify and interpret intricate spatial patterns and relationships.

This section introduces various DL models suitable for analyzing ST data, including graph neural networks (GNNs), CNNs, transformer models, and hybrid approaches (see Fig. 4).

This figure highlights the key DL models used in ST, including GNNs, CNNs, transformer models, and hybrid approaches, showcasing their unique processing techniques and applications in cancer research.

3.1.1 Graph-based deep learning models—Graph-based models excel at contextualizing the complex morphology and structure within whole slide images by representing spatial relationships through graph structures. These models efficiently encode the intricate patterns of tissue architecture, enabling a detailed analysis of ST data. This approach facilitates the exploration of cellular interactions and molecular pathways across diverse tissue regions, enhancing our understanding of disease mechanisms and progression (Wu et al., 2022; Yuan & Bar-Joseph, 2020). Azher et al. (2023) demonstrated the potential of leveraging ST data with a contrastive crossmodal pretraining mechanism to improve graph-based learning tasks. This approach enhanced the extraction of molecular and histological information, which improved outcomes in cancer staging, lymph node metastasis prediction, survival prediction, and tissue clustering analyses. The integration of spatial omics data significantly improved the performance of graph-based models in pathology workflows, thus highlighting their potential to enhance cancer diagnostics and prognostics.

Peng, He, Peng, Li, and Zhang (2023a) developed STGNNks, a method that combines GNNs, denoising auto-encoders, and k-sums clustering to process spatially resolved transcriptomics data. This model constructs a hybrid adjacency matrix and integrates gene expressions with spatial context, mapping learned features to a low-dimensional space for clustering. STGNNks significantly outperformed existing ST analysis algorithms and provided valuable insights into tumor progression through spatial trajectory inference and differentially expressed gene detection. Additionally, Wu et al. (2022) applied GNNs to model TMEs as local subgraphs using spatial protein profiles. This strategy captures distinctive cellular interactions associated with differential clinical outcomes. The model demonstrated substantial improvements in predicting patient outcomes compared to traditional DL approaches, underscoring its potential to enhance the understanding of TME dynamics.

3.1.2 Convolutional neural networks (CNNs)—CNNs are adept at analyzing histopathological images and predicting gene expression patterns due to their ability to extract high-level features from complex visual data. These networks use multiple layers of filters to process spatial hierarchies in images and capture details from cell morphology to tissue architecture (Dabeer, Khan, & Islam, 2019). This enables CNNs to identify subtle patterns and structures that are crucial for accurate gene expression prediction,

which enhances our understanding of the molecular underpinnings of cancer and other diseases. Zheng, Carrillo-Perez, Pizurica, Heiland, and Gevaert (2023) developed a DL model to predict transcriptional subtypes of glioblastoma cells from histology images. This model phenotypically analyzed millions of tissue spots and identified consistent associations between tumor architecture and prognosis across independent cohorts. The approach underscored the critical connection between spatial cellular architecture and clinical outcomes, thus offering a scalable method to predict gene expression and understand the spatial organization of tumor cells.

In a related advancement, Wang, Zhou, Kong, and Lu (2023) proposed a CNN-based method for enhancing spot resolution in ST, termed superresolved ST. Their approach, utilizing a shift-predict operation, achieved 9× superresolution and outperformed traditional superresolution techniques. This method provides a deeper understanding of gene expression patterns and their underlying biological significance, highlighting the potential of CNNs to revolutionize ST by improving spatial resolution and analytical precision.

3.1.3 Transformer models—Transformers excel in capturing long-range dependencies, which enhances gene expression predictions from histology images by leveraging the self-attention mechanism introduced in the seminal paper *Attention Is All You Need* (Vaswani et al., 2023). This approach allows transformers to simultaneously process distant parts of an image, identifying crucial patterns and correlations for accurate prediction. This mechanism enables the comprehensive analysis of the whole tissue section, improving the mapping of gene expressions and aiding in the study of complex interactions within the TME. Liu, Zhang, and Luo (2024) proposed a contrastive learning-based framework to infer molecular subtypes and clinical outcomes of breast cancer from unannotated whole slide images. By leveraging patch-level features and a gated attention mechanism, the model produced slide-level predictions. This method effectively established high-order genotype-phenotype associations, enhancing the application of digital pathology in clinical practice. Xiao, Kong, Li, Wang, and Lu (2024) introduced TCGN, a model combining convolutional layers, transformer encoders, and GNNs to estimate gene expressions from histopathological images. TCGN operates with a single spot image input, maintaining interpretability while enhancing accuracy. The model provided a powerful tool for inferring gene expressions in precision health applications.

The versatility of transformers to capture long-range dependencies complements the localized feature extraction of CNNs. Combining these approaches can further enhance model performance, leading us to hybrid approaches.

3.1.4 Hybrid approaches—Hybrid models combine multiple DL architectures to enhance the analysis of ST data. By integrating the unique advantages of each model, they achieve a more comprehensive and accurate interpretation of complex biological data. This synergy allows for a deeper understanding of the spatial and molecular dynamics within tissues, aiding in the identification of key biological insights and potential therapeutic targets (Song et al., 2024; Yan et al., 2019).

Steyaert et al. (2023a) emphasized the importance of developing multimodal fusion approaches to integrate complementary data types such as molecular, histopathology, radiology, and clinical records. By leveraging DL models that can handle multiple modalities, researchers can capture the heterogeneity of complex diseases more accurately, advancing precision medicine. Chen, Zhang, Tang, Liu, and Huang (2024) proposed the Edge-relational Window-attentional GNN (ErwaNet) for predicting gene expression from standard tissue images. This model constructs heterogeneous graphs to model local window interactions and incorporates an attention mechanism for global information analysis. ErwaNet stands out as a cost-effective and accurate method for gene expression prediction, offering a significant advantage in cancer research by providing more efficient and accessible analytical paradigms.

By integrating the strengths of different models, hybrid approaches offer a robust framework for ST analysis, combining the localized feature extraction of CNNs, the contextual understanding of transformers, and the relational mapping of GNNs.

3.2 Successful applications of deep learning in multimodal spatial transcriptomics

This subsection showcases how DL significantly enhances the analysis of ST by enabling precise cell-type identification, biomarker discovery, and treatment response prediction. Each case study directly correlates with the models discussed, illustrating the practical impact of these technological advancements.

Table 3 lists recent DL applications in ST, emphasizing diverse methodologies and their impacts on cancer research. Each entry will be succinctly described, highlighting its relevance and connection to the DL models introduced earlier.

These case studies illustrate the diverse applications of DL in enhancing ST analysis. By integrating multimodal data, such as histological images and ST, DL models have significantly improved our understanding of cellular and molecular mechanisms in various diseases.

Reflecting on this comprehensive review, some key areas highlight the significant strides where DL optimizes multimodal ST:

DL methods, such as STGNNks and RESEPT, have revolutionized cell-type identification and spatial domain mapping in tissue samples. By precisely characterizing cellular heterogeneity and tissue organization, these advancements open new avenues for understanding disease mechanisms and developing targeted therapies tailored to specific cell populations (Chang et al., 2022; Peng et al., 2023a). Additionally, DL approaches, exemplified by the work of Halawani et al. (2023) and Liu et al. (2024), have accelerated biomarker discovery by analyzing spatial gene expression patterns and tumor heterogeneity. The identification of novel molecular markers through these techniques offers the potential for earlier disease detection, more accurate prognosis, and the development of targeted therapies that address the unique molecular signatures of individual tumors.

Furthermore, DL models, as demonstrated by Wu et al. (2022) and Zheng et al. (2023), have shown remarkable promise in predicting patient outcomes and treatment responses. By

analyzing spatial relationships and cellular interactions within tumors, these models offer a powerful tool for tailoring treatment strategies to individual patients, ultimately leading to improved clinical outcomes and more effective cancer care.

These advancements underscore the transformative potential of DL in ST, which is paving the way for more precise and personalized medical interventions. The integration of multimodal data, supported by robust DL frameworks is continuing to drive significant progress in cancer research and precision medicine. By leveraging these advanced techniques, researchers can significantly improve the accuracy, robustness, and interpretability of ST analysis, ultimately accelerating the development of novel diagnostics and therapeutics in precision medicine.

4. Impact of multimodal spatial transcriptomics on cancer diagnostics and therapeutics

In the preceding sections, we explored how multimodal ST integrates various data types to revolutionize cancer research. This section will focus on how these integrations enhance cancer diagnostics and therapeutics, thus highlighting specific instances where ST has uniquely contributed to advancements in these areas. Here, we emphasize novel applications and methodologies that demonstrate the transformative potential of ST in clinical settings.

4.1 Enhancing cancer diagnosis

Recent advancements have demonstrated the power of pathogenomics in integrating advanced molecular diagnostics from genomic data with morphological information from histopathological imaging and codified clinical data. Feng et al. (2024) describe how this integration leads to the discovery of new multimodal cancer biomarkers, and thus enhances the precision of oncology diagnoses. Their study emphasizes the importance of synthesizing complementary modalities with emerging multimodal AI methods in pathogenomics, including the correlation and fusion of histology and genomics profiles of cancer. This approach overcomes the limitations of unimodal data and propels precision oncology into the next decade.

Similarly, Xi et al. (2022) presented a modality-correlation embedding model for breast tumor diagnosis using mammography and ultrasound images. By optimizing the correlation between these modalities, their model improved tumor classification accuracy significantly, with a sensitivity of 91.67% and a specificity of 95.83%. This study demonstrates the importance of maintaining pairwise closeness of multimodal data in a common label space with consistent diagnostic results from multimodal images of the same patient.

Abdollahyan et al. (2023) highlight the role of dynamic biobanking in advancing breast cancer research. The Breast Cancer Now Tissue Bank at the Barts Cancer Institute serves as a dynamic biobanking ecosystem and integrates longitudinal biospecimens with multimodal data, including EHRs, genomic, and imaging data. This ecosystem supports precision medicine efforts in breast cancer research by offering high-quality annotated biospecimens and advanced analytics tools, demonstrating how integrated data can inform diagnostic accuracy.

4.2 Personalized treatment strategies

Chen et al. (2023) introduced the Multimodal Data Fusion Diagnosis Network (MDFNet), a framework that effectively fuses clinical skin images with patient clinical data for skin cancer classification. MDFNet establishes a mapping among heterogeneous data features which addresses issues of feature paucity and richness that arise from using single-mode data. Their model showed an improvement in diagnostic accuracy by about 9% and illustrates the fusion advantages exhibited by MDFNet and its potential as an effective auxiliary diagnostic tool for skin cancer, enhancing clinical decision-making.

Yan et al. (2019) developed a hybrid DL method for breast cancer classification by integrating pathological images and structured data from clinical EMRs. This approach significantly outperforms state-of-the-art methods and demonstrates the importance of combining multimodal data for precise patient stratification and personalized therapy. Their method provides a practical tool for breast cancer diagnosis in real clinical practice.

Rani, Ahmad, Masood, and Saxena (2023) highlighted the importance of using machine learning models on unimodal and multimodal datasets for diagnosing breast cancer molecular subtypes. By operating on multimodal data samples for each patient, their custom DL-based model pipeline achieved 94% accuracy and thus showed the superiority of multimodal data over unimodal in breast cancer subtype classification. These findings underscore the value of multimodal data in enhancing the precision of treatment strategies.

Li et al. (2022) utilized multi-omics data for lung cancer stage prediction, showing that microbial combinatorial transcriptome fusion analysis had the highest accuracy, reaching 0.809. This study reveals the role of multimodal data and fusion algorithms in accurately diagnosing lung cancer stages, which aids doctors in clinics and emphasizes the potential of multimodal integration in personalized treatment planning.

4.3 Deep learning’s contribution to ST-based precision oncology

Steyaert et al. (2023a) discussed the challenges and opportunities in multimodal data fusion for cancer biomarker discovery with DL. They highlighted the need for developing effective multimodal fusion approaches as a single modality might not be sufficient to capture the heterogeneity of complex diseases like cancer. By tackling data sparsity and scarcity, enhancing multimodal interpretability, and standardizing datasets, DL can significantly advance the analysis of biomedical data to the discovery of key cancer biomarkers.

Lobato-Delgado, Priego-Torres, and Sanchez-Morillo (2022) combined molecular, imaging, and clinical data analysis for predicting cancer prognosis and demonstrated how integration of these data types using machine learning and DL techniques improves patient stratification and clinical management. This review emphasizes that multimodal models can better stratify patients and can contribute to personalized medicine and provide valuable insights into cancer biology and progression.

Zhang et al. (2024b) explored the use of AI in liver imaging, summarizing AI methodologies and their applications in managing liver diseases. They stressed the importance of feature

interpretability, multimodal data integration, and multicenter study in enhancing the clinical applications of AI, supporting precise disease detection, diagnosis, and treatment planning.

Shao, Ma, Zhang, Li, and Wang (2023) discussed predicting gene mutation status via AI technologies based on multimodal integration to advance precision oncology. Their review synthesized the general framework of MMI for molecular intelligent diagnostics and summarized emerging applications of AI in predicting mutational profiles of common cancers. This work highlights the potential of AI as a decision-support tool to aid oncologists in future cancer treatment management and the challenges and prospects of AI in medical fields.

As we conclude our exploration of the impact that multimodal ST and DL have on enhancing cancer diagnostics and therapeutics, we recognize the transformative advancements these technologies bring to precision oncology. The integration of diverse data types, coupled with sophisticated analytical models, has not only revolutionized our approach to understanding and treating cancer but also set the stage for addressing the complexities of implementation in clinical settings.

Looking ahead, Section 5 will delve into the challenges, future outlook, and ethical considerations of these technologies. We will explore the barriers that must be overcome to fully realize their potential, the ethical frameworks that need to be established for their responsible use, and the directions future research might take to continue advancing the field of cancer research and treatment.

5. Challenges, future outlook, and ethical considerations

As we delve into the transformative world of AI-driven multimodal ST in cancer research and treatment, it’s critical to navigate the complex challenges and ethical considerations while also looking toward future advancements. By addressing these elements comprehensively, we can fully harness the transformative potential of these technologies.

5.1 Challenges and future directions

5.1.1 Challenges in multimodal spatial transcriptomics—While traditional ST has revolutionized our understanding of tissue architecture and cellular function, it faces several issues that hinder its broader application. These include data sparsity and noise, potential batch effects, and missing data, all of which can obscure meaningful biological insights. Despite significant advancements, the field of multimodal ST continues to encounter numerous challenges that must be addressed to fully unlock its transformative potential. Integrating and interpreting high-dimensional data from diverse sources requires sophisticated computational tools and algorithms. Moreover, the heterogeneity and sparsity of single-cell and ST data add layers of complexity to the analysis.

A major barrier in the field is the lack of standardization and interoperability among different imaging and omics modalities. Standardized protocols and interdisciplinary collaboration are essential to overcome these challenges (Jaiswal, Agarwal, & Jaiswal,

2023). Robust imputation techniques and model-based approaches are critical for addressing missing and sparse data, which are pivotal for maintaining the integrity of ST studies (Rathore, Abdulkadir, & Davatzikos, 2020). Technical variability and batch effects can significantly skew analyses, necessitating robust preprocessing pipelines and advanced batch correction methods to mitigate these effects (Feng et al., 2024; Khalighi et al., 2024). The scalability of data analysis and the ethical, legal, and social implications must be addressed for the responsible implementation of AI and ST in healthcare (Khalighi et al., 2024; Liu et al., 2024).

Standardization and robust imputation techniques are crucial for combating issues of missing data and ensuring interoperability across platforms (Capobianco, 2022; Rathore et al., 2020). These measures will help maintain data integrity and enhance the comparability of results across different studies. Addressing these challenges will pave the way for the more effective and widespread use of multimodal ST in research and clinical settings, ultimately advancing our understanding of complex biological systems and improving patient care.

5.1.2 Future directions—To overcome the existing challenges and pave the way for new discoveries, research in multimodal ST is focusing on several key areas:

Increasing sequencing depth and using robust preprocessing pipelines are necessary strategies to enhance the signal-to-noise ratio and address data sparsity issues (Akhoundova & Rubin, 2022; Feng et al., 2024). Implementing advanced batch correction methods and developing multimodal fusion techniques are essential to mitigate batch effects and improve data integration (Khalighi et al., 2024; Steyaert et al., 2023a). Ensuring the reliability and generalizability of models through practices like cross-validation and external validation is crucial. Additionally, incorporating interpretability techniques, such as attention mechanisms and SHAP values, helps translate complex ST data into actionable insights (Steyaert et al., 2023a; Yu et al., 2023). Rigorous biological validation of model predictions is essential to ensure that identified patterns reflect true biological phenomena. Integrating ST data with experimental data, such as functional assays, is crucial for confirming the biological significance of the findings.

Promoting collaborative efforts and establishing standardized workflows for data sharing are key to advancing ST research and developing standardized analytical tools (Tortora et al., 2023). Continued advancements in computational models, including DL and Bayesian statistics, will enhance data analysis capabilities and offer new insights into complex biological systems (Li, 2023). Building comprehensive databases and repositories for multimodal data will facilitate data sharing and collaborative research, which are crucial for accelerating discoveries in cancer research. Integrating clinical data with multimodal ST is vital for translating research into clinical practice and developing more effective and personalized therapies.

By addressing these challenges and leveraging future directions, the field of multimodal ST is poised for significant advancements, promising to enhance our understanding of complex biological systems and improve therapeutic strategies in precision medicine.

5.2 Ethical dimensions and responsible AI use

The use of AI and ST in healthcare introduces significant ethical considerations that must be managed to ensure responsible application:

5.2.1 Ethical considerations in data collection and analysis—Ensuring that patient data is collected and used in a manner that respects consent and confidentiality is paramount to balancing the benefits of AI-driven diagnostics against potential risks to patient autonomy and privacy (Love, 2023). Building trust among patients and healthcare providers requires transparent AI algorithms that provide clear explanations of their decision-making processes, ensuring that decisions are made in the best interests of patients. Ethical guidelines for data anonymization and sharing should be established to protect patient identities and outline the purposes for which data can be used, fostering responsible research practices.

5.2.2 Addressing biases and ensuring fairness—Training AI algorithms on diverse datasets is essential to prevent biases that could lead to healthcare disparities (Wang et al., 2024). Ensuring that AI models provide equitable care and identifying any biases that could disadvantage certain patient groups are crucial for fostering fair healthcare practices (Fusar-Poli et al., 2022). Developing comprehensive ethical frameworks to address the challenges and risks associated with AI in healthcare is necessary to guide ethical implementation and ensure model accuracy and fairness (Sundaramurthy & Vaithiyalingam, 2023).

In summary, addressing the challenges in ST data analysis demands a multifaceted approach that includes improving data quality, developing robust validation strategies, enhancing model interpretability, and performing rigorous biological validation. By tackling these challenges head-on, researchers can fully leverage the potential of ST technologies to revolutionize precision medicine and deepen our understanding of complex biological systems. By addressing these challenges and ethical considerations, we can enhance the robustness, interpretability, and application of ST in cancer research. Continued development and integration of these advanced tools promise to transform our understanding of cancer, improve diagnostic and therapeutic strategies, and ultimately enhance patient outcomes.

6. Conclusion

The integration of multimodal ST and DL has already begun to transform cancer research, offering deep insights into tumor biology and enabling personalized treatment strategies. As we reflect on the journey through this chapter, the diverse applications and potential of multimodal ST have been thoroughly explored, emphasizing its critical role in enhancing diagnostic precision, guiding personalized therapies, and advancing our understanding of cancer at the molecular level.

Key takeaways revisited.

- **Enhanced diagnostic accuracy:** The integration of various data types through multimodal ST has proven to improve diagnostic accuracy and precision, highlighting the importance of comprehensive views of tumor biology.
- **Personalized treatment strategies:** By utilizing the detailed molecular and spatial information provided by ST, researchers have developed targeted therapeutic strategies, enhancing personalized medicine approaches.
- **DL and AI integration:** DL models, such as CNNs and GNNs, have been essential in analyzing complex ST data, supporting the advancement of precision oncology.
- **Ethical considerations and responsible AI use:** Ensuring ethical practices in AI and ST integration remains crucial for maintaining transparency, fairness, and patient data protection in clinical settings.

Transformative potential of multimodal ST and DL.

The advancements in multimodal ST and DL are not merely incremental; they are transformative, enabling a holistic understanding of cancer. These technologies facilitate the development of more effective and less toxic treatments by providing insights into the TME and the interactions of different cell types and molecular pathways within the spatial context of tissues.

As we look to the future, the promise of these technologies to revolutionize cancer research, improve diagnostic and therapeutic strategies, and ultimately enhance patient outcomes is immense. Addressing the remaining ethical and technical challenges will be crucial to fully realize the potential of these cutting-edge tools in cancer research and clinical practice.

In conclusion, the journey through multimodal ST and DL in cancer research has been enlightening, demonstrating the profound impact these technologies will continue to have on advancing our understanding of cancer and improving patient care. As we continue to develop and integrate these tools, the future of cancer research looks brighter than ever.

References

Abdollahyan M, Gadaleta E, Asif M, Oscanoa J, Barrow-McGee R, Jones S, et al. (2023). Dynamic biobanking for advancing breast cancer research. Journal of Personalized Medicine, 13(2), 360. 10.3390/jpm13020360. [PubMed: 36836597]

Acosta JN, Falcone GJ, Rajpurkar P, & Topol EJ (2022). Multimodal biomedical AI. Nature Medicine, 28(9), 1773–1784. 10.1038/s41591-022-01981-2.

Akhoundova D, & Rubin MA (2022). Clinical application of advanced multi-omics tumor profiling: Shaping precision oncology of the future. Cancer Cell, 40(9), 920–938. 10.1016/j.ccell.2022.08.011. [PubMed: 36055231]

Aneja S, Aneja N, Abas PE, & Naim AG (2021). Transfer learning for cancer diagnosis in histopathological images. arXiv preprint, arXiv:2112.15523. 10.48550/arXiv.2112.15523.

Azher Z, Fatemi M, Lu Y, Srinivasan G, Diallo A, Christensen B, et al. (2023). Spatial omics driven crossmodal pretraining applied to graph-based deep learning for cancer pathology analysis. bioRxiv: The Preprint Server for Biology, 2023.07.30.551187. 10.1101/2023.07.30.551187.

Boehm KM, Khosravi P, Vanguri R, Gao J, & Shah SP (2022). Harnessing multimodal data integration to advance precision oncology. Nature Reviews. Cancer, 22(2), 114–126. 10.1038/s41568-021-00408-3. [PubMed: 34663944]

Bottosso M, Mosele F, Michiels S, Cournède P-H, Dogan S, Labaki C, et al. (2024). Moving toward precision medicine to predict drug sensitivity in patients with metastatic breast cancer. ESMO Open, 9(3), 102247. 10.1016/j.esmoop.2024.102247. [PubMed: 38401248]

Cai M, Zhao L, Hou G, Zhang Y, Wu W, Jia L, et al. (2023). FDTrans: Frequency domain transformer model for predicting subtypes of lung cancer using multimodal data. Computers in Biology and Medicine, 158, 106812. 10.1016/j.compbiomed.2023.106812. [PubMed: 37004434]

Capobianco E (2022). High-dimensional role of AI and machine learning in cancer research. British Journal of Cancer, 126(4), 523–532. 10.1038/s41416-021-01689-z. [PubMed: 35013580]

Cascarano A, Mur-Petit J, Hernández-González J, Camacho M, de Toro Eadie N, Gkontra P, et al. (2023). Machine and deep learning for longitudinal biomedical data: A review of methods and applications. Artificial Intelligence Review, 56(2), 1711–1771. 10.1007/s10462-023-10561-w.

Cebula H, Noel G, Garnon J, Todeschi J, Burckel H, de Mathelin M, et al. (2020). The cryo-immunologic effect: A therapeutic advance in the treatment of glioblastomas? Neurochirurgie, 66(6), 455–460. 10.1016/j.neuchi.2020.06.135. [PubMed: 33045247]

Chang Y, He F, Wang J, Chen S, Li J, Liu J, et al. (2022). Define and visualize pathological architectures of human tissues from spatially resolved transcriptomics using deep learning. Computational and Structural Biotechnology Journal, 20, 4600–4617. 10.1016/j.csbj.2022.08.029. [PubMed: 36090815]

Chen C, Zhang Z, Tang P, Liu X, & Huang B (2024). Edge-relational window-attentional graph neural network for gene expression prediction in spatial transcriptomics analysis. Computers in Biology and Medicine, 174, 108449. 10.1016/j.compbiomed.2024.108449. [PubMed: 38626512]

Chen Q, Li M, Chen C, Zhou P, Lv X, & Chen C (2023). MDFNet: Application of multimodal fusion method based on skin image and clinical data to skin cancer classification. Journal of Cancer Research and Clinical Oncology, 149(7), 3287–3299. 10.1007/s00432-022-04180-1. [PubMed: 35918465]

Cristian A, Rubens M, Orada R, DeVries K, Syrkin G, DePiero MT, et al. (2024). Development of a cancer rehabilitation dashboard to collect data on physical function in cancer patients and survivors. American Journal of Physical Medicine & Rehabilitation, 103(3S Suppl 1), S36–S40. 10.1097/PHM.0000000000002424. [PubMed: 38364028]

Dabeer S, Khan MM, & Islam S (2019). Cancer diagnosis in histopathological image: CNN based approach. Informatics in Medicine Unlocked, 16, 100231. 10.1016/j.imu.2019.100231.

Dabral I, Singh M, & Kumar K (2021). Cancer detection using convolutional neural network. In Tripathi M, & Upadhyaya S (Eds.). Conference Proceedings of ICDLAIR2019 (pp. 290–298)Springer International Publishing. 10.1007/978-3-030-67187-7_30.

Dawood M, Eastwood M, Jahanifar M, Young L, Ben-Hur A, Branson K, et al. (2023). Cross-linking breast tumor transcriptomic states and tissue histology. Cell Reports. Medicine, 4(12), 101313. 10.1016/j.xcrm.2023.101313. [PubMed: 38118424]

Deng G, Yang J, Zhang Q, Xiao Z-X, & Cai H (2018). MethCNA: A database for integrating genomic and epigenomic data in human cancer. BMC Genomics, 19(1), 138. 10.1186/s12864-018-4525-0. [PubMed: 29433427]

Du J, Yang Y-C, An Z-J, Zhang M-H, Fu X-H, Huang Z-F, et al. (2023). Advances in spatial transcriptomics and related data analysis strategies. Journal of Translational Medicine, 21, 330. 10.1186/s12967-023-04150-2. [PubMed: 37202762]

Emmert-Buck MR, Bonner RF, Smith PD, Chuaqui RF, Zhuang Z, Goldstein SR, et al. (1996). Laser capture microdissection. Science (New York, N.Y.), 274(5289), 998–1001. 10.1126/science.274.5289.998. [PubMed: 8875945]

Feng X, Shu W, Li M, Li J, Xu J, & He M (2024). Pathogenomics for accurate diagnosis, treatment, prognosis of oncology: A cutting edge overview. Journal of Translational Medicine, 22(1), 131. 10.1186/s12967-024-04915-3. [PubMed: 38310237]

Fusar-Poli P, Manchia M, Koutsouleris N, Leslie D, Woopen C, Calkins ME, et al. (2022). Ethical considerations for precision psychiatry: A roadmap for research and clinical practice. European Neuropsychopharmacology, 63, 17–34. 10.1016/j.euroneuro.2022.08.001. [PubMed: 36041245]

Halawani R, Buchert M, & Chen Y-PP (2023). Deep learning exploration of single-cell and spatially resolved cancer transcriptomics to unravel tumour heterogeneity. Computers in Biology and Medicine, 164, 107274. 10.1016/j.compbiomed.2023.107274. [PubMed: 37506451]

Hanczar B, Bourgeais V, & Zehraoui F (2022). Assessment of deep learning and transfer learning for cancer prediction based on gene expression data. BMC Bioinformatics, 23(1), 262. 10.1186/s12859-022-04807-7. [PubMed: 35786378]

Harikumar H, Quinn TP, Rana S, Gupta S, & Venkatesh S (2021). Personalized single-cell networks: A framework to predict the response of any gene to any drug for any patient. BioData Mining, 14(1), 37. 10.1186/s13040-021-00263-w. [PubMed: 34353329]

Hu B, Sajid M, Lv R, Liu L, & Sun C (2022). A review of spatial profiling technologies for characterizing the tumor microenvironment in immuno-oncology. Frontiers in Immunology, 13, 996721. 10.3389/fimmu.2022.996721. [PubMed: 36389765]

Jaiswal H, Agarwal A, & Jaiswal V (2023). Expanding the scope of oncological understanding via multimodal imaging. Onkologia i Radioterapia, 17(10), 652–661.

Jia Y, Liu J, Chen L, Zhao T, & Wang Y (2024). THItoGene: A deep learning method for predicting spatial transcriptomics from histological images. Briefings in Bioinformatics, 25(1), bbad464. 10.1093/bib/bbad464.

Jiang X, Hu Z, Wang S, & Zhang Y (2023). Deep learning for medical image-based cancer diagnosis. Cancers, 15(14), 3608. 10.3390/cancers15143608. [PubMed: 37509272]

Johnson KB, Wei W, Weeraratne D, Frisse ME, Misulis K, Rhee K, et al. (2021). Precision medicine, AI, and the future of personalized health care. Clinical and Translational Science, 14(1), 86–93. 10.1111/cts.12884. [PubMed: 32961010]

Joo S, Ko ES, Kwon S, Jeon E, Jung H, Kim J-Y, et al. (2021). Multimodal deep learning models for the prediction of pathologic response to neoadjuvant chemotherapy in breast cancer. Scientific Reports, 11(1), 18800. 10.1038/s41598-021-98408-8. [PubMed: 34552163]

Kattau M, Glocker B, & Darambara D (2021). A comparative analysis of two deep learning architectures for the automatic segmentation of vestibular schwannoma. IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2021, 1–3. 10.1109/NSS/MIC44867.2021.9875733.

Khalighi S, Reddy K, Midya A, Pandav KB, Madabhushi A, & Abedalthagafi M (2024). Artificial intelligence in neuro-oncology: Advances and challenges in brain tumor diagnosis, prognosis, and precision treatment. NPJ Precision Oncology, 8(1), 1–12. 10.1038/s41698-024-00575-0. [PubMed: 38167869]

Li J, Li L, You P, Wei Y, & Xu B (2023). Towards artificial intelligence to multi-omics characterization of tumor heterogeneity in esophageal cancer. Seminars in Cancer Biology, 91, 35–49. 10.1016/j.semcancer.2023.02.009. [PubMed: 36868394]

Li M, Jiang Y, Zhang Y, & Zhu H (2023). Medical image analysis using deep learning algorithms. Frontiers in Public Health, 11, 1273253. 10.3389/fpubh.2023.1273253. [PubMed: 38026291]

Li Q (2023). AI-powered Bayesian statistics in biomedicine. Statistics in Biosciences, 15(3), 737–749. 10.1007/s12561-023-09400-x.

Li W, Liu B, Wang W, Sun C, Che J, Yuan X, et al. (2022). Lung cancer stage prediction using multi-omics data. Computational and Mathematical Methods in Medicine, 2022, 2279044. 10.1155/2022/2279044. [PubMed: 35880092]

Liang X, Young WC, Hung L-H, Raftery AE, & Yeung KY (2019). Integration of multiple data sources for gene network inference using genetic perturbation data. Journal of Computational Biology, 26(10), 1113–1129. 10.1089/cmb.2019.0036. [PubMed: 31009236]

Lipkova J, Chen RJ, Chen B, Lu MY, Barbieri M, Shao D, et al. (2022). Artificial intelligence for multimodal data integration in oncology. Cancer Cell, 40(10), 1095–1110. 10.1016/j.ccell.2022.09.012. [PubMed: 36220072]

Liu H, Zhang Y, & Luo J (2024). Contrastive learning-based histopathological features infer molecular subtypes and clinical outcomes of breast cancer from unannotated whole slide

images. Computers in Biology and Medicine, 170, 107997. 10.1016/j.compbiomed.2024.107997. [PubMed: 38271839]

Liu Y, Wu Z, Feng Y, Gao J, Wang B, Lian C, et al. (2023). Integration analysis of single-cell and spatial transcriptomics reveal the cellular heterogeneity landscape in glioblastoma and establish a polygenic risk model. Frontiers in Oncology, 13, 1109037. 10.3389/fonc.2023.1109037. [PubMed: 37397378]

Lobato-Delgado B, Priego-Torres B, & Sanchez-Morillo D (2022). Combining molecular, imaging, and clinical data analysis for predicting cancer prognosis. Cancers, 14(13), 3215. 10.3390/cancers14133215. [PubMed: 35804988]

Love CS (2023). "Just the facts ma'am": Moral and ethical considerations for artificial intelligence in medicine and its potential to impact patient autonomy and hope. The Linacre Quarterly, 90(4), 375–394. 10.1177/00243639231162431. [PubMed: 37974568]

Luo R, & Bocklitz T (2023). A systematic study of transfer learning for colorectal cancer detection. Informatics in Medicine Unlocked, 40, 101292. 10.1016/j.imu.2023.101292.

Lyubetskaya A, Rabe B, Fisher A, Lewin A, Neuhaus I, Brett C, et al. (2022). Assessment of spatial transcriptomics for oncology discovery. Cell Reports Methods, 2(11), 100340. 10.1016/j.crmeth.2022.100340. [PubMed: 36452860]

Mayampurath A, Ramesh S, Michael D, Liu L, Feinberg N, Granger M, et al. (2021). Predicting response to chemotherapy in patients with newly diagnosed high-risk neuroblastoma: A report from the International Neuroblastoma Risk Group (Scopus) JCO Clinical Cancer Informatics, 5, 1181–1188. 10.1200/CCI.21.00103. [PubMed: 34882497]

Patton RM, Johnston JT, Young SR, Schuman CD, Potok TE, Rose DC, . & Saltz J (2019). Exascale deep learning to accelerate cancer research. In 2019 IEEE International Conference on Big Data (Big Data) (pp. 1488–1496). IEEE. 10.1109/BigData47090.2019.9006467.

Peng L, He X, Peng X, Li Z, & Zhang L (2023a). STGNNks: Identifying cell types in spatial transcriptomics data based on graph neural network, denoising auto-encoder, and k-sums clustering. Computers in Biology and Medicine, 166, 107440. 10.1016/j.compbiomed.2023.107440. [PubMed: 37738898]

Peng Z, Ren Z, Tong Z, Zhu Y, Zhu Y, & Hu K (2023b). Interactions between MFAP5 + fibroblasts and tumor-infiltrating myeloid cells shape the malignant microenvironment of colorectal cancer. Journal of Translational Medicine, 21(1), 405. 10.1186/s12967-023-04281-6. [PubMed: 37344903]

Pham TD (2023). Prediction of five-year survival rate for rectal cancer using Markov models of convolutional features of RhoB expression on tissue microarray. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 20(5), 3195–3204. 10.1109/TCBB.2023.3274211. [PubMed: 37155403]

Price K, Westfall AD, Beemer C, Casper LH, Dierkes D, Hickey EM, et al. (2022). Building and sustaining a comprehensive pediatric oncology care team: The roles and integration of psychosocial and rehabilitative team members. Horizons in Cancer Research, 84, 1–63.

Rahman MN, Noman AA, Turza AM, Abrar MA, Samee MAH, & Rahman MS (2023). ScribbleDom: Using scribble-annotated histology images to identify domains in spatial transcriptomics data. Bioinformatics (Oxford, England), 39(10), btad594. 10.1093/bioinformatics/btad594.

Rani S, Ahmad T, Masood S, & Saxena C (2023). Diagnosis of breast cancer molecular subtypes using machine learning models on unimodal and multimodal datasets. Neural Computing and Applications, 35(34), 24109–24121. 10.1007/s00521-023-09005-x.

Rathore S, Abdulkadir A, & Davatzikos C (2020). Analysis of MRI data in diagnostic neuroradiology. Annual Review of Biomedical Data Science, 3(1), 365–390. 10.1146/annurev-biodatasci-022620-015538.

Schmitt M, Sinnberg T, Niessner H, Forschner A, Garbe C, Macek B, et al. (2021). Individualized proteogenomics reveals the mutational landscape of melanoma patients in response to immunotherapy. Cancers, 13(21), 5411. 10.3390/cancers13215411. [PubMed: 34771574]

Shao J, Ma J, Zhang Q, Li W, & Wang C (2023). Predicting gene mutation status via artificial intelligence technologies based on multimodal integration (MMI) to advance precision oncology. Seminars in Cancer Biology, 91, 1–15. 10.1016/j.semcancer.2023.02.006. [PubMed: 36801447]

Sharma P, Nayak DR, Balabantaray BK, Tanveer M, & Nayak R (2024). A survey on cancer detection via convolutional neural networks: Current challenges and future directions. Neural Networks, 169, 637–659. 10.1016/j.neunet.2023.11.006. [PubMed: 37972509]

Song J, Lamstein J, Ramaswamy VG, Webb M, Zada G, Finkbeiner S, et al. (2024). Enhancing spatial transcriptomics analysis by integrating image-aware deep learning methods. Pacific Symposium on Biocomputing, 29, 450–463. [PubMed: 38160299]

Song T, Markham KK, Li Z, Muller KE, Greenham K, & Kuang R (2022). Detecting spatially co-expressed gene clusters with functional coherence by graph-regularized convolutional neural network. Bioinformatics (Oxford, England), 38(5), 1344–1352. 10.1093/bioinformatics/btab812. [PubMed: 34864909]

Ståhl PL, Salmén F, Vickovic S, Lundmark A, Navarro JF, Magnusson J, et al. (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science (New York, N. Y.), 353(6294), 78–82. 10.1126/science.aaf2403. [PubMed: 27365449]

Steyaert S, Pizurica M, Nagaraj D, Khandelwal P, Hernandez-Boussard T, Gentles AJ, et al. (2023a). Multimodal data fusion for cancer biomarker discovery with deep learning. Nature Machine Intelligence, 5(4), 351–362. 10.1038/s42256-023-00633-5.

Steyaert S, Qiu YL, Zheng Y, Mukherjee P, Vogel H, & Gevaert O (2023b). Multimodal deep learning to predict prognosis in adult and pediatric brain tumors. Communications Medicine, 3(1), 1–15. 10.1038/s43856-023-00276-y. [PubMed: 36596859]

Sundaramurthy A, & Vaithiyalingam C (2023). Ethical dimensions and future prospects of artificial intelligence in decision making systems for oncology: A comprehensive analysis and reference scheme. In: 2023 International conference on intelligent technologies for sustainable electric and communications systems (iTech SECOM) (pp. 59–63). 10.1109/iTechSECOM59882.2023.10435323.

Tortora M, Cordelli E, Sicilia R, Nibid L, Ippolito E, Perrone G, et al. (2023). RadioPathomics: Multimodal learning in non-small cell lung cancer for adaptive radiotherapy. IEEE Access, 11, 47563–47578. 10.1109/ACCESS.2023.3275126.

Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, & Polosukhin I (2023). Attention is all you need. arXiv preprint, arXiv:1706.03762. 10.48550/arXiv.1706.03762.

Wang J, Li B, Luo M, Huang J, Zhang K, Zheng S, et al. (2024). Progression from ductal carcinoma in situ to invasive breast cancer: Molecular features and clinical significance. Signal Transduction and Targeted Therapy, 9(1), 1–28. 10.1038/s41392-024-01779-3. [PubMed: 38161204]

Wang S, Zhou X, Kong Y, & Lu H (2023). Superresolved spatial transcriptomics transferred from a histological context. Applied Intelligence, 53(24), 31033–31045. 10.1007/s10489-023-05190-3.

Wei Z, Zhang Y, Weng W, Chen J, & Cai H (2021). Survey and comparative assessments of computational multi-omics integrative methods with multiple regulatory networks identifying distinct tumor compositions across pan-cancer data sets. Briefings in Bioinformatics, 22(3), bbaa102. 10.1093/bib/bbaa102.

Williams CG, Lee HJ, Asatsuma T, Vento-Tormo R, & Haque A (2022). An introduction to spatial transcriptomics for biomedical research. Genome Medicine, 14(1), 68. 10.1186/s13073-022-01075-1. [PubMed: 35761361]

Wilson CM, Li K, Yu X, Kuan P-F, & Wang X (2019). Multiple-kernel learning for genomic data mining and prediction. BMC Bioinformatics, 20(1), 426. 10.1186/s12859-019-2992-1. [PubMed: 31416413]

Wu Z, Trevino AE, Wu E, Swanson K, Kim HJ, D'Angio HB, et al. (2022). Graph deep learning for the characterization of tumour microenvironments from spatial protein profiles in tissue specimens. Nature Biomedical Engineering, 6(12), 1435–1448. 10.1038/s41551-022-00951-w.

Xi X, Li W, Li B, Li D, Tian C, & Zhang G (2022). Modality-correlation embedding model for breast tumor diagnosis with mammography and ultrasound images. Computers in Biology and Medicine, 150, 106130. 10.1016/j.compbiomed.2022.106130. [PubMed: 36215846]

Xiao X, Kong Y, Li R, Wang Z, & Lu H (2024). Transformer with convolution and graph-node co-embedding: An accurate and interpretable vision backbone for predicting gene expressions from local histopathological image. Medical Image Analysis, 91, 103040. 10.1016/j.media.2023.103040. [PubMed: 38007979]

Xu C, Jin X, Wei S, Wang P, Luo M, Xu Z, et al. (2022). DeepST: Identifying spatial domains in spatial transcriptomics by deep learning. Nucleic Acids Research, 50(22), e131. 10.1093/nar/gkac901. [PubMed: 36250636]

Yan R, Ren F, Rao X, Shi B, Xiang M, … Zhang L, et al. (2019). Integration of multimodal data for breast cancer classification using a hybrid deep learning method. Intelligent computing theories and application. ICIC 2019. Lecture notes in computer science. Cham: Springer460–469. 10.1007/978-3-030-26763-6_44.

Yu X, Liu R, Gao W, Wang X, & Zhang Y (2023). Single-cell omics traces the heterogeneity of prostate cancer cells and the tumor microenvironment. Cellular & Molecular Biology Letters, 28(1), 38. 10.1186/s11658-023-00450-z. [PubMed: 37161356]

Yuan Y, & Bar-Joseph Z (2020). GCNG: Graph convolutional networks for inferring gene interaction from spatial transcriptomics data. Genome Biology, 21(1), 300. 10.1186/s13059-020-02214-w. [PubMed: 33303016]

Zhang N, Yang X, Piao M, Xun Z, Wang Y, Ning C, et al. (2024a). Biomarkers and prognostic factors of PD-1/PD-L1 inhibitor-based therapy in patients with advanced hepatocellular carcinoma. Biomarker Research, 12(1), 26. 10.1186/s40364-023-00535-z. [PubMed: 38355603]

Zhang P, Gao C, Huang Y, Chen X, Pan Z, Wang L, et al. (2024b). Artificial intelligence in liver imaging: Methods and applications. Hepatology International, 18(2), 422–434. 10.1007/s12072-023-10630-w. [PubMed: 38376649]

Zhen X, Chen J, Zhong Z, Hrycushko B, Zhou L, Jiang S, et al. (2017). Deep convolutional neural network with transfer learning for rectum toxicity prediction in cervical cancer radiotherapy: A feasibility study. Physics in Medicine and Biology, 62(21), 8246–8263. 10.1088/1361-6560/aa8d09. [PubMed: 28914611]

Zheng Y, Carrillo-Perez F, Pizurica M, Heiland DH, & Gevaert O (2023). Spatial cellular architecture predicts prognosis in glioblastoma. Nature Communications, 14(1), 4122. 10.1038/s41467-023-39933-0.

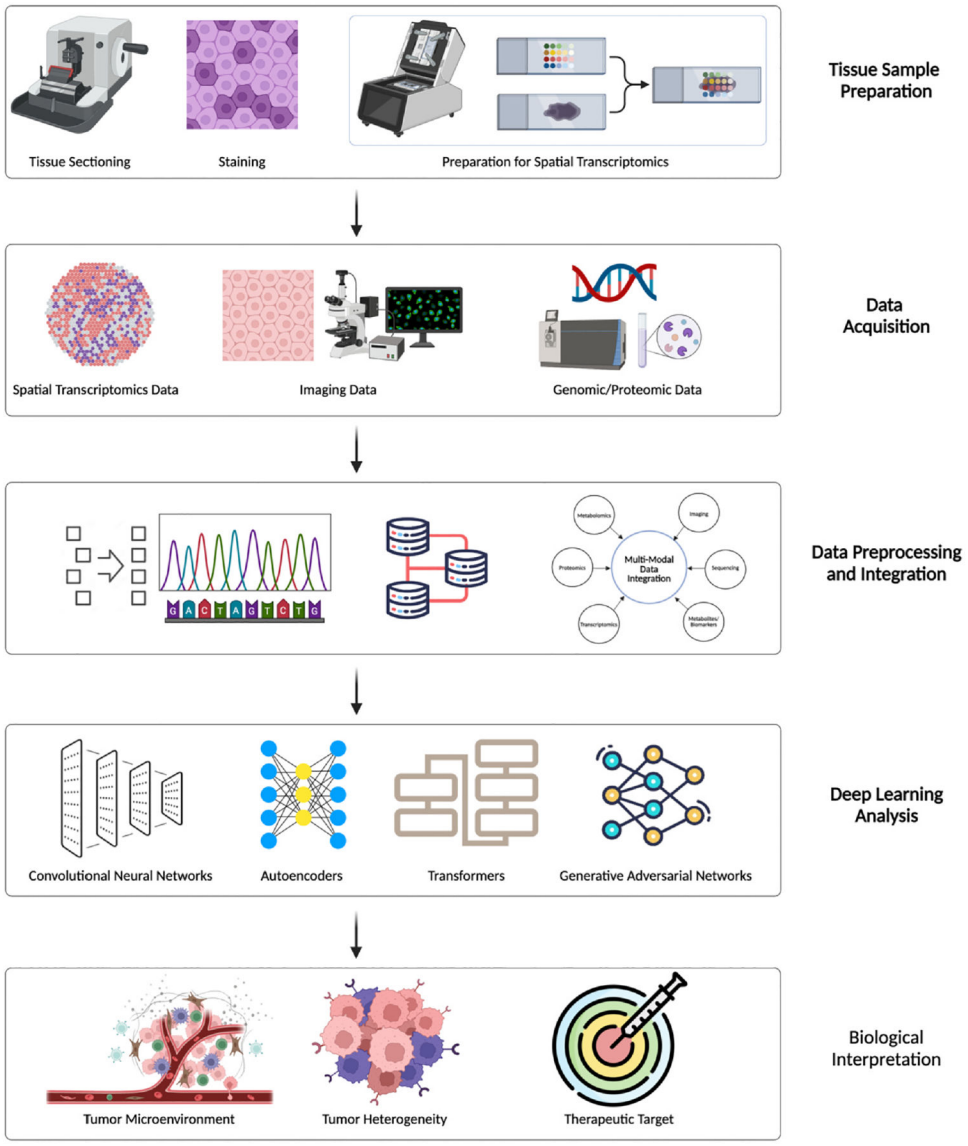

Fig. 1.

Deep learning-based multimodal spatial transcriptomics workflow for cancer analysis.

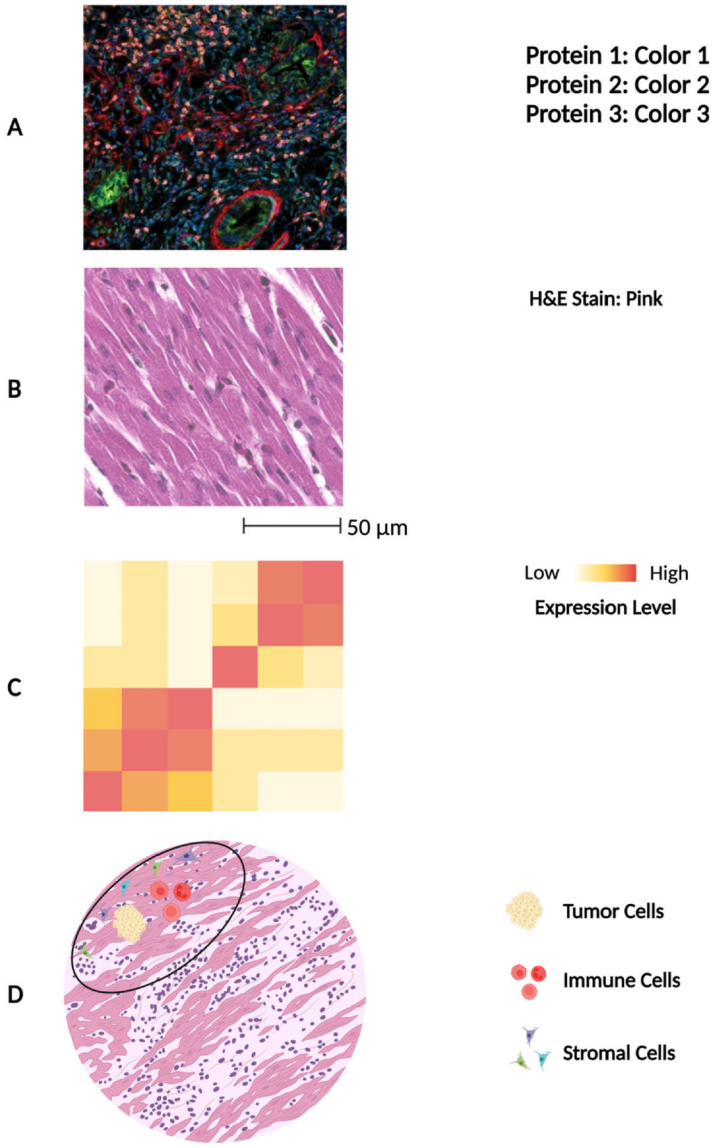

Fig. 2.
Visualization of multimodal spatial transcriptomics data in a cancer tissue section.

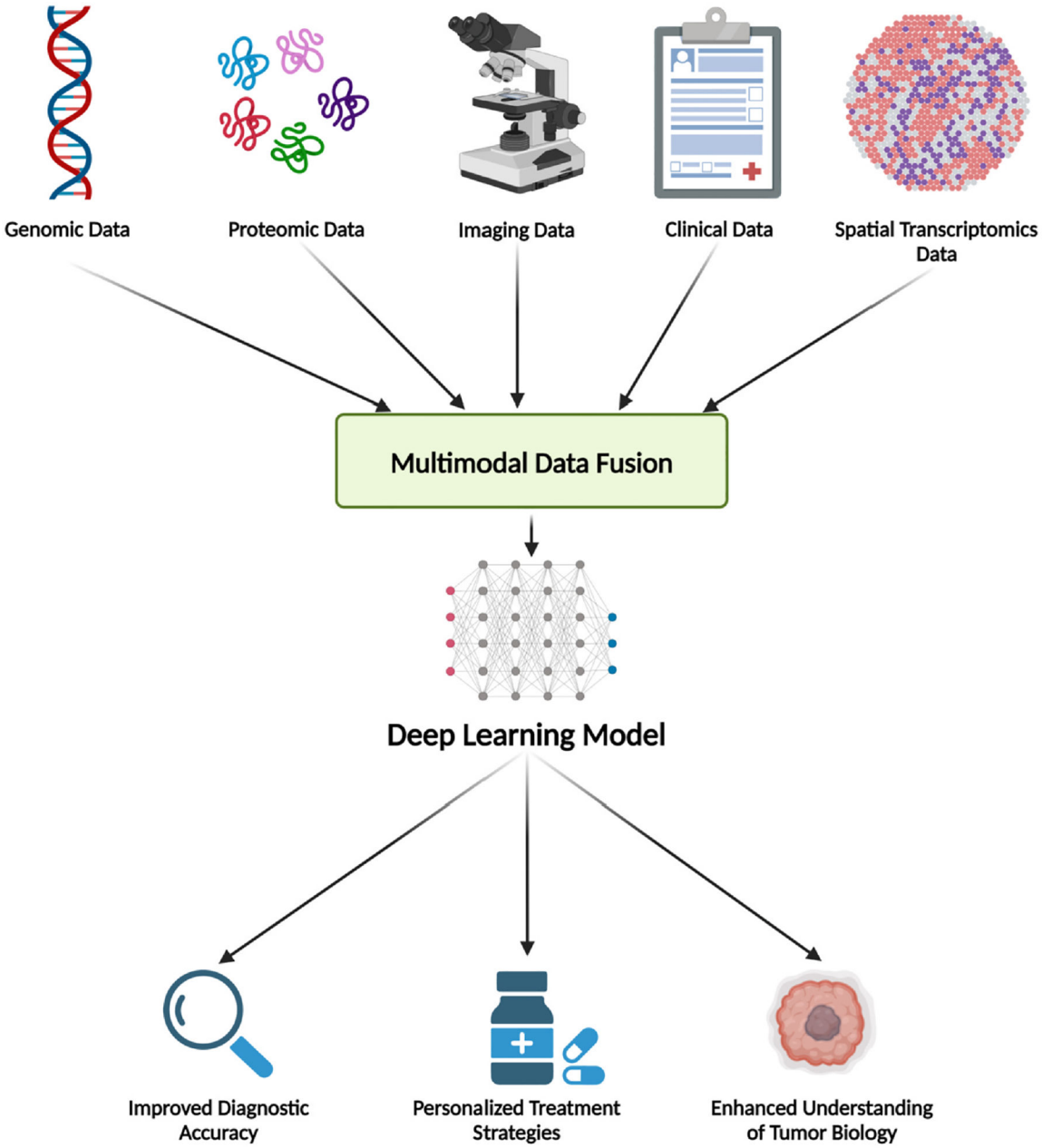

Fig. 3.

Integration of genomic, proteomic, imaging, clinical, and spatial transcriptomics data using deep learning models in cancer research.

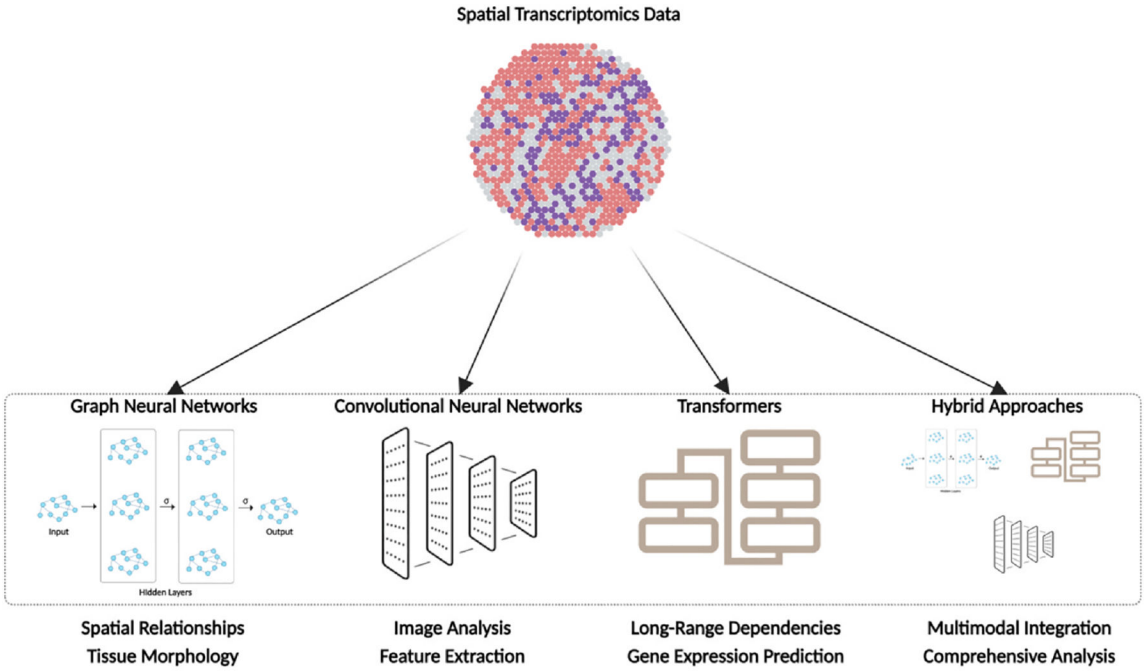

Fig. 4.
Different deep learning models for analyzing spatial transcriptomics data.

Table 1

Studies on deep learning in oncology.

Title	Key insights
Cancer Detection Using Convolutional Neural Network (Dabral, Singh, & Kumar, 2021)	Developed a CNN-based method for classifying cancer cells into benign and malignant, enhancing early cancer detection accuracy
Deep convolutional neural network with transfer learning for rectum toxicity prediction in cervical cancer radiotherapy: a feasibility study (Zhen et al., 2017)	Used a pre-trained convolutional neural network (VGG-16 CNN) and transfer learning to predict rectum toxicity in cervical cancer radiotherapy, showing significant opportunities for deep learning in radiation oncology
Exascale deep learning to accelerate cancer research (Patton et al., 2019)	Demonstrated that neural network architectures could be automatically generated and tailored for specific applications, significantly accelerating cancer pathology research
A comparative analysis of two deep learning architectures for the automatic segmentation of vestibular schwannoma (Kattau, Glocker, & Darambara, 2021)	Investigated the performance of two existing deep learning frameworks for tumor segmentation, highlighting advancements in automatic segmentation techniques for treatment planning
Prediction of five-year survival rate for rectal cancer using Markov models of convolutional features of RhoB expression on tissue microarray (Pham, 2023)	Combined deep learning, data coding, and probabilistic modeling to predict five-year survival rates in rectal cancer patients, achieving higher accuracy compared to traditional methods

Table 2

Examples of multimodal data integration in cancer research.

Title	Key insights
Building and sustaining a comprehensive pediatric oncology care team: The roles and integration of psychosocial and rehabilitative team members (Price et al., 2022)	Highlighted the importance of integrating psychosocial and rehabilitative care services in pediatric oncology to improve patient outcomes and quality of life
The cryo-immunologic effect: A therapeutic advance in the treatment of glioblastomas? (Cebula et al., 2020)	Discussed the potential of combining cryotherapy with immunotherapy to enhance the immune response against glioblastomas, demonstrating the benefits of a multimodal therapeutic approach
Predicting response to chemotherapy in patients with newly diagnosed high-risk neuroblastoma (Mayampurath et al., 2021)	Utilized convolutional neural networks and clinical models to predict chemotherapy response in high-risk neuroblastoma patients, highlighting the effectiveness of integrating imaging and clinical data for treatment planning
Development of a cancer rehabilitation dashboard to collect data on physical function in cancer patients and survivors (Cristian et al., 2024)	Described the creation of a dashboard that integrates self-reported and objective physical function data, facilitating personalized rehabilitation plans and improving patient management in oncology
RadioPathomics: Multimodal learning in non-small cell lung cancer for adaptive radiotherapy (Tortora et al., 2023)	Developed a multimodal late fusion approach combining radiomics, pathomics, and clinical data to predict radiotherapy outcomes in non-small-cell lung cancer patients, demonstrating improved precision in treatment planning through data integration

Table 3

Recent applications of deep learning in multimodal spatial transcriptomics.

Title	Key applications	Methodology	Key findings	Impact
Spatial omics driven crossmodal pretraining applied to graph-based deep learning for cancer pathology analysis (Azher et al., 2023)	Cancer Pathology Analysis	Graph Neural Network, Crossmodal Pretraining	Enhances extraction of molecular and histological information, improving cancer staging and survival prediction	Improved pathology workflow accuracy
STGNNks: Identifying cell types in spatial transcriptomics data based on graph neural network, denoising auto-encoder, and k-sums clustering (Peng et al., 2023a)	Cell Type Identification	Graph Neural Network, Auto-encoder, K-sums Clustering	Outperforms other clustering methods, providing insights into tumor progression and spatial gene expression	Superior spatial transcriptomics data analysis and disease diagnosis
Graph deep learning for the characterization of tumor microenvironments from spatial protein profiles in tissue specimens (Wu et al., 2022)	Tumor Microenvironment Analysis	Graph Neural Network	Identifies spatial motifs associated with cancer recurrence and patient survival, outperforming traditional models	Enhanced understanding of tumor microenvironments
Single-cell omics traces the heterogeneity of prostate cancer cells and the tumor microenvironment (Yu, Liu, Gao, Wang, & Zhang, 2023)	Tumor Heterogeneity Analysis	Single-cell Omics, Deep Learning	Analyzes heterogeneity and spatial distribution of prostate cancer cells, revealing new insights into tumor biology	Advanced prostate cancer research and personalized treatment strategies
Spatial cellular architecture predicts prognosis in glioblastoma (Zheng et al., 2023)	Prognosis Prediction	Deep Learning, Histology Analysis	Links tumor architecture with prognosis, identifying gene expression patterns associated with survival.	Improved prognosis prediction for glioblastoma patients
Contrastive learning-based histopathological features infer molecular subtypes and clinical outcomes of breast cancer from unannotated whole slide images (Liu et al., 2024)	Molecular Subtype Prediction	Contrastive Learning, Attention Mechanism	Predicts molecular subtypes and clinical outcomes, establishing genotype-phenotype associations	Enhanced molecular subtype prediction for breast cancer
Transformer with convolution and graph-node co-embedding: An accurate and interpretable vision backbone for predicting gene expressions from local histopathological image (Xiao et al., 2024)	Gene Expression Prediction	Transformer, Convolutional Layers, Graph Neural Networks	Operates with single spot image input, maintaining interpretability while enhancing accuracy	Improved gene expression prediction from histopathological images
Detecting spatially co-expressed gene clusters with functional coherence by graph-regularized convolutional neural network (Song et al., 2022)	Gene Cluster Detection	Convolutional Neural Network, Graph Regularization	Effectively detects spatially co-expressed gene clusters, providing biological insights into tissue morphology	Enhanced understanding of tissue morphology and gene functionality
Deep learning exploration of single-cell and spatially resolved cancer transcriptomics to unravel tumour heterogeneity (Halawani et al., 2023)	Tumor Heterogeneity Analysis	Deep Learning Frameworks	Enhances precision oncology applications by analyzing single-cell and spatial transcriptomics data to investigate tumor heterogeneity	Advanced tumor characterization and personalized oncology
Superresolved spatial transcriptomics transferred from a histological context (Wang et al., 2023)	Spatial Resolution Enhancement	Convolutional Neural Network	Achieves superresolution in spatial transcriptomics data, improving gene expression pattern analysis	Enhanced spatial resolution and gene expression analysis
THItoGene: a deep learning method for predicting spatial transcriptomics from histological images (Jia, Liu, Chen, Zhao, & Wang, 2024)	Gene Expression Prediction	Dynamic Convolutional and Capsule Networks	Demonstrates superior performance in spatial gene expression prediction, linking pathology image phenotypes to gene expression regulation	Improved link between pathology images and gene expression
Define and visualize pathological architectures of human tissues from spatially resolved transcriptomics using deep learning (Chang et al., 2022)	Tissue Architecture Visualization	Deep Learning Framework (RESEPT)	Accurately infers and visualizes tissue architecture, aiding in the identification of pathological features in diseases like Alzheimer's and cancer	Better visualization and understanding of tissue architecture

Title	Key applications	Methodology	Key findings	Impact
ScribbleDom: Using scribble-annotated histology images to identify domains in spatial transcriptomics data (Rahman et al., 2023)	Domain Identification	Semi-Supervised Convolutional Neural Network	Improves spatial domain identification by integrating human annotations with computational power	Enhanced spatial domain identification and integration of human expertise
Cross-linking breast tumor transcriptomic states and tissue histology (Dawood et al., 2023)	Gene Expression Prediction	Graph Neural Network	Captures gene expression states from whole slide images, linking histological phenotypes with gene expression patterns	Better linking of histological and gene expression data
DeepST: identifying spatial domains in spatial transcriptomics by deep learning (Xu et al., 2022)	Spatial Domain Identification	Deep Learning Framework	Provides accurate and scalable spatial domain identification, outperforming existing methods	More accurate spatial domain identification