



MVASA-HGN: multi-view adaptive semantic-aware heterogeneous graph network for *KRAS* mutation status prediction

Wanting Yang^{1^}, Shinichi YOSHIDA², Juanjuan Zhao^{1,3,4}, Wei Wu⁵, Yan Qiang^{1,6}

¹College of Computer Science and Technology (College of Data Science), Taiyuan University of Technology, Taiyuan, China; ²School of Informatics, Kochi University of Technology, Kochi, Japan; ³College of Software, Taiyuan University of Technology, Taiyuan, China; ⁴Jinzhong College of Information, Taiyuan, China; ⁵Department of Clinical Laboratory, Shanxi Provincial People's Hospital, Taiyuan, China; ⁶School of Software, North University of China, Taiyuan, China

Contributions: (I) Conception and design: W Yang; (II) Administrative support: S YOSHIDA, J Zhao, Y Qiang; (III) Provision of study materials or patients: W Wu; (IV) Collection and assembly of data: J Zhao, W Wu; (V) Data analysis and interpretation: W Yang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Yan Qiang, PhD. College of Computer Science and Technology (College of Data Science), Taiyuan University of Technology, No. 79 West Street Yingze, Taiyuan 030024, China; School of Software, North University of China, No. 3 Xueyuan Road, Taiyuan 030051, China. Email: qiangyan@tyut.edu.cn.

Background: In the treatment of advanced non-small cell lung cancer (NSCLC), the mutation status of the Kirsten rat sarcoma virus oncogene homolog (*KRAS*) gene has been shown to be a key factor affecting the efficacy of immune checkpoint inhibitors (ICIs), which is an important guideline for physicians to develop personalized treatment strategies. However, existing mutation prediction studies have primarily focused on the feature representation of individual patient medical data, ignoring the complex semantic relationships among patients in diverse clinical features. This study aimed to accurately identify *KRAS* gene status, which will not only assist physicians in accurately screening the patient population most likely to benefit from immunotherapy, but also reduce patient burden by avoiding unnecessary treatment attempts.

Methods: A multi-view adaptive semantics-aware heterogeneous graph framework (MVASA-HGN) based on multimodal medical data was developed to accurately predict *KRAS* mutation status in NSCLC patients. The framework first parses the relational semantics through clinical feature clustering and constructs a heterogeneous graph by combining computed tomography (CT) image and clinical features. In the second step, the heterogeneous graph is split into relational subgraphs under multiple views, and the node representations are constructed and updated gradually through a two-stage strategy of single-view graph representation learning and multi-view heterogeneous information fusion. In the single-view phase, we enhance the node self-embedding and construct the adjacency embedding of neighbors with the same type of relationship to ensure that the relational subgraph under each semantic preserves the complete local structure. Two attention mechanisms are introduced in the multi-view fusion phase to capture the enriched semantics preserved in nodes and heterogeneous relations, respectively. Finally, a comprehensive node representation is obtained through adaptive aggregation of different view neighborhood information and enhanced node embedding without predefined meta-paths.

Results: The classification results were evaluated on cooperative hospitals and The Cancer Imaging Archive (TCIA) datasets, and ablation experiments and comparison experiments were performed on the components of the framework, while exploring the framework's rationality and interpretability. Accuracy reached 85.29% and specificity reached 89.67% on the test set, indicating that our framework has significant advantages in deeply modeling complex heterogeneous semantics in local structures and fully exploiting

[^] ORCID: 0000-0002-5896-027X.

and utilizing the rich semantic information preserved in heterogeneous relationships. The source code of MVASA-HGN is available at <https://github.com/Yangwanter37/MVASA-HGN>.

Conclusions: Our proposed MVASA-HGN framework provides a new perspective for multimodal information fusion and creates a new avenue to explore the potential link between images and genes, and the framework provides a non-invasive and cost-effective solution for identifying *KRAS* mutation status, which has a broad application prospect.

Keywords: *KRAS* mutation prediction; heterogeneous graph neural networks (GNNs); multi-view learning

Submitted Jul 05, 2024. Accepted for publication Nov 20, 2024. Published online Jan 21, 2025.

doi: 10.21037/qims-24-1370

View this article at: <https://dx.doi.org/10.21037/qims-24-1370>

Introduction

As one of the most prevalent malignant tumors worldwide, lung cancer, with its persistently high morbidity and mortality rates, remains an urgent challenge in global health (1,2). Despite modern medicine's innovation in treatment strategies, the 5-year survival rate of lung cancer patients is still not optimistic (3). More notably, there are significant differences in treatment response and prognostic outcomes among patients with the same type of lung cancer, highlighting the importance of personalized treatment strategies. In light of this, a deeper exploration of the molecular biology of lung cancer, especially the mutation status of oncogenic driver genes, has become crucial to enhance the therapeutic efficacy (4). Among them, Kirsten rat sarcoma virus oncogene homolog (*KRAS*), as one of the most commonly mutated genes in non-small cell lung cancer (NSCLC) (5,6), can activate the RAS/MAPK signaling pathway through the conversion of its amino acid residue at position 12 from glycine (G) to cysteine (C) (7), which then promotes the proliferation and metastasis of tumor cells (8). Therefore, accurate detection of *KRAS* gene mutation status is significant in guiding the development of personalized lung cancer treatment regimens.

In recent years, immune checkpoint inhibitors (ICIs) have opened new avenues for NSCLC treatment (9). Studies have shown that tumors carrying *KRAS* mutations exhibit higher sensitivity to ICIs due to their higher tumor mutational burden (TMB) (10) and abundance of tumor-infiltrating lymphocytes (TILs) (11), providing a solid theoretical basis for immunotherapy (12). Specifically, the KEYNOTE-042 study showed (13) that pembrolizumab significantly prolonged progression-free survival (PFS) in patients with *KRAS* mutations and was significantly more efficacious than non-mutated patients (15 *vs.* 6 months),

emphasizing the critical guiding role of precise detection of *KRAS* mutation status in the development of treatment strategies. This strategy enables more effective screening of patient populations that respond better to ICIs while avoiding unnecessary treatments and their potential toxic responses (4,14). Although biopsy tissue sequencing is the gold standard for genetic testing, its limitations are the difficulty of obtaining adequate specimens for some patients and the fact that biopsies may increase the risk of cancer metastasis under certain circumstances (15,16). Therefore, the development of a novel, noninvasive, and easily accessible method to accurately identify the *KRAS* gene mutation status in NSCLC patients is essential to promote the process of personalization of lung cancer treatment.

Computed tomography (CT), as a noninvasive diagnostic technique, plays a crucial role in routine clinical practice tasks in lung cancer, such as screening, staging, assessing response to treatment, and monitoring recurrence, owing to its ability to capture rich pathophysiologic information (17,18). The correlation between lung CT image features and gene mutation status has recently received extensive attention (19,20). For example, Nair *et al.* (21) developed a multivariate logistic regression model based on texture features in CT and fluorodeoxyglucose (FDG) positron emission tomography (PET)-CT images of NSCLC patients, which effectively predicted the mutation status of epidermal growth factor receptor (*EGFR*). Meanwhile, Wang *et al.* (22) proposed an innovative adaptive model that significantly improves the accuracy of *EGFR* mutation prediction by sequentially learning local and global features from different lobes of the lungs through a segmentation network and using a domain adaptive strategy to learn robust features in CT images of different thicknesses. In addition, the decision tree algorithm model designed

by Luo *et al.* (23) and the random forest model designed by Jia *et al.* (24) both achieved noninvasive prediction of the *EGFR* mutation status of lung adenocarcinoma based on the radiomics features of CT images. The study by Morgado *et al.* (25) extended the field of view to the entire lung region of interest, explored the relationship between image phenotype and *EGFR* mutation status deeply, and achieved superior prediction performance compared with localized nodal analysis. For multi-task learning research on gene mutation prediction, Moreno *et al.* (26) proposed a new selective category average voting (SCAV) scheme with an integrated approach to improve the performance of *EGFR* and *KRAS* mutation prediction in small sample data. In contrast, Sun *et al.* (27) introduced deep learning (DL) to combine multi-task learning to simultaneously predict *EGFR* and *KRAS* mutation status using the ResNet network and attention mechanism, demonstrating the great potential of DL in predicting gene mutation status. Furthermore, Zhang *et al.* (28) innovatively combined serum tumor markers and CT imaging features to improve the accuracy of *EGFR* mutation prediction by logistic regression prediction model, which provided more comprehensive information support for clinical decision-making.

However, medical data is inherently multimodal, and this heterogeneity provides patients with complementary diagnostic perspectives. Multiple forms of medical data complement each other and can provide physicians with multidimensional information that helps them make more accurate clinical decisions. Existing studies have shown that combining imaging features with patients' clinical characteristics can enhance the model's performance to a certain extent (29–32). For example, Yang *et al.* (29) constructed a nomogram model that can be applied to different types of CT based on radiomics features and clinical features, which can accurately identify the *EGFR* mutation status of NSCLC patients. Its prediction performance was significantly improved after combining clinical features, demonstrating multimodal data fusion's advantages. Similarly, Gao *et al.* (30) extracted radiomics features from PET/CT images of lung adenocarcinoma patients and constructed nine joint radiomics models by combining clinical parameters. The experimental results showed that combining clinical parameters could significantly improve the performance of predicting *EGFR* mutation status. Kim *et al.* (31) used a DL model to predict *EGFR* mutations by combining radiological features, deep features, and clinical data from pre-treatment CTs of NSCLC patients. Recently, Chen *et al.* (32) first proposed

an *EGFR* prediction model based on stacked DL, which integrates information extracted from PET/CT and clinical data and shows the powerful prediction capability of stacked DL. Although these studies have achieved great results in combining imaging and clinical features to predict gene mutation status, these methods are often limited to simple feature splicing and analyzed only from an individual perspective without considering the deeper interconnections between patients and the correlations or complementary relationships between different modalities. Therefore, how to effectively integrate multimodal medical data into a unified framework to fully capture the interconnections between patients in this context has become an urgent and challenging problem to be solved.

In recent years, graph neural networks (GNNs) have made significant progress in capturing complex associations. Compared with convolutional neural networks (CNNs), GNNs focus more on local feature aggregation of graph topology, which provides greater flexibility in parsing deep relationships between heterogeneous features in cross-modal data (33). In order to deal with diverse relationship types in graph networks, multi-view graph networks provide a solution idea that has been explored in the medical field. For example, Zhang *et al.* (34) proposed a multi-view GNN, which combines information from functional connectivity (FC) and resting-state functional magnetic resonance imaging (rs-fMRI). They constructed a multi-view model using self-attention graph pooling and graph CNNs to identify major depressive disorder (MDD) effectively. Al-Sabri *et al.* (35) developed an automatic multi-view GNN framework designed to automatically extract biomedical entities and relationships, which captures the embedding of multi-relationship nodes through automatic multi-view representation learning, which in turn enhances the representation capability of the graph network. In addition, Xiao *et al.* (36) designed an end-to-end learning framework in combination with comparative learning to effectively integrate multiple prior knowledge into a GNN for analyzing multi-omics data deeply. Fan *et al.* (37) then combined five biological graph features and multi-omics data to predict individual cancer cells' synthetic lethality (SL) using a multi-view graph convolutional network (GCN) model. They integrated graph-specific representations obtained from GCN models through a max pooling layer and accurately predicted SL interactions using a deep neural network. Meanwhile, Zhang *et al.* (38) constructed a multi-view multimodal feature fusion network by considering the multidimensional nature of clinical

information in prognosis prediction. This network enhanced the fine-grained feature interactions between patient time series and clinical information through a co-attention module, then constructed a patient correlation graph using structural information from clinical records, and finally fused the multimodal features of patients using a GNN. However, despite the progress of these studies in processing multiple types of associations in multi-view graph networks, most of these studies have mainly focused on biological entity and brain image analysis, which is difficult to directly apply to CT images to predict *KRAS* gene mutation status in NSCLC patients. In addition, these studies still need to delve further into the heterogeneity of relationship types between different views, especially in modelling node attribute semantics and heterogeneous relationship semantics, to obtain a more comprehensive node embedding.

And in contrast to homogeneous GNNs, heterogeneous graphs can more realistically show the structural associations between data and rich semantic information by introducing different types of nodes and edges. In order to accurately capture this heterogeneous information, current heterogeneous graph representation learning methods generally rely on meta-paths to extract critical semantic information. For example, Han *et al.* (39) proposed a multi-relational knowledge graph embedding representation learning method by introducing an attention mechanism. They utilized meta-path guiding to form heterogeneous awareness nodes, decomposed the knowledge graph embedding into two parts: structural embedding and multi-relational embedding, and efficiently characterized the entities and relations through joint learning, thus improving the accuracy of knowledge reasoning. However, in medical scenarios, relying on meta-paths alone may not fully reflect the complex patient relationships. To overcome this challenge, Yang *et al.* (40) proposed an intelligent diagnostic model based on a heterogeneous GCN, which focuses on the intrinsic attributes of the symptoms as well as the multiple hidden relationships between symptoms and attributes. However, this approach is still limited by predefined meta-paths, which restricts the model's generalization ability. Dai *et al.* (41) further designed a multi-relational graph attention network (GAT), significantly improving the model's performance by adapting the importance of different neighboring nodes through the self-attention layer. Peng *et al.* (42) proposed a flexible neighbor selection-guided multi-relational GNN architecture with a label-aware similarity metric and an enhanced relation-aware neighbor selection mechanism,

improving the model's generalization ability and robustness. Baek *et al.* (43) took a different perspective, in which they considered that patients with similar illnesses may be treated similarly. Based on this point of view, they used GNNs to mine multi-contextual information to predict emerging health risks, providing personalized health risk assessments for patients with chronic diseases. However, these studies are still deficient in fusing multimodal information, which is especially critical in medical research because patient data usually contains multiple modalities from different sources and forms. Therefore, how to effectively fuse multimodal information in heterogeneous graph structures has become an urgent direction to be explored. To this end, D'souza *et al.* (44) proposed a novel multiplex GNN, which tracks the information flow in a multiplex graph through a message-passing walking system to effectively capture the complex cross-modal dependencies among features, providing the necessary flexibility for mining multimodal data. In addition, Jia *et al.* (45) also proposed a new multimodal heterogeneous GAT, which adaptively captures the heterogeneity information of the graph through edge-level aggregation and uses a modality-level attention mechanism to obtain multimodal fusion information. This further improves the model's ability to process multimodal data and overall performance. In summary, although the heterogeneous graph network approach has been shown to be beneficial, there are still some issues that need further study on a deeper level: how to reduce the dependence on domain knowledge? How to efficiently pass multiple semantic information of patients? How to distinguish the importance of neighboring nodes under different relationship semantics during information aggregation?

Therefore, this study proposes a graph framework for heterogeneous semantic-aware in multiple views (MVSA-HGN) to predict *KRAS* gene mutation status in NSCLC patients. The framework constructs heterogeneous GNNs under multiple views based on multimodal medical data to comprehensively model semantic information in nodes and relationships. Specifically, we divide different views based on the clustering results of clinical features to effectively model the local structure under different heterogeneous relationships. Given the heterogeneity of relationship types among different views, we design two complementary attention mechanisms to capture node attribute semantics and heterogeneous relationship semantics, respectively, and adaptively fuse the node's own enhanced embedding to obtain the final representation of the node on this basis. The experimental results fully demonstrate the effectiveness

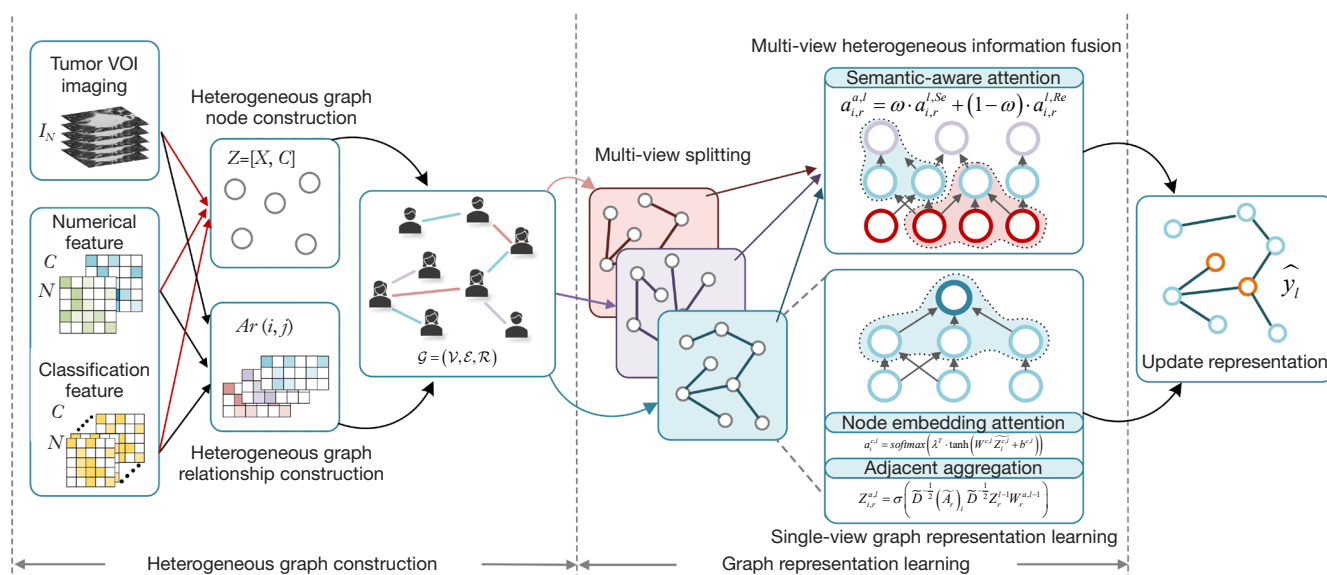


Figure 1 Overview of the MVASA-HGN. Firstly, image features are extracted while clustering clinical features to construct a heterogeneous graph network using their similarity, and then split it into multiple views based on heterogeneous semantics. Within each view, enhance the node's own embedding and construct the neighbor embedding of homogeneous relation types. Then in the multi-view fusion stage, two attention mechanisms are introduced to capture the semantics preserved in the nodes and heterogeneous relations, respectively, to adaptively aggregate the information from different views and the enhanced node embeddings to obtain comprehensive node representations and realize the prediction of gene states. VOI, volume of interest; MVASA-HGN, multi-view adaptive semantics-aware heterogeneous graph network.

of MVASA-HGN.

The main contributions of this work are as follows:

- (I) An MVASA-HGN modeling inter-patient relationships is constructed for predicting the *KRAS* gene mutation status of NSCLC patients. This framework can effectively fuse information from multimodal medical data to provide personalized medication guidance for patients.
- (II) A multi-level joint attention mechanism is proposed to jointly model multi-view heterogeneous information from two perspectives: node attribute semantics and heterogeneous relationship semantics, and adaptively inductively learn the discriminative features in the complex associations between patients, thus avoiding over-reliance on predefined meta-paths and domain knowledge, improving the generalization ability of the model.
- (III) The model's effectiveness was fully evaluated in both the collaborative hospital and the public datasets. The experimental results show that our model outperforms existing methods and models complex associations between patients from a novel perspective, which helps to identify underlying

disease patterns and influencing factors, providing a powerful practical tool for personalized lung cancer treatment.

Methods

Overview

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by Shanxi Provincial People's Hospital Ethics Committee (No. 2023.299) and the requirement for individual consent for this retrospective analysis was waived. This paper proposes an MVASA-HGN approach, for which the procedure is shown in *Figure 1*. The method aims to analyze the complex node and relationship semantics in heterogeneous graphs from a multi-view perspective and contains three core steps. First, image features are extracted by a pre-trained encoder, and clinical information is semantically grouped using clustering methods. The heterogeneous graph is constructed based on these images and clinical features. Second, the heterogeneous graph is split into relational subgraphs under multiple views

based on the relational semantics, and a view-specific adjacency representation is generated in each view by aggregating the information of neighboring nodes. Finally, the inherent associations between nodes and relational semantics are explored profoundly while preserving the complete semantic information. Based on obtaining enhanced node embeddings, complex heterogeneous knowledge, and semantic information from different views are adaptively fused to construct comprehensive and rich node representations for the target nodes. Finally, the fused embeddings are fed into the fully connected layer to yield the final classification results.

Heterogeneous graph construction

Problem definition

This study involves two types of heterogeneous data: structured data (patient age, gender, smoking history, etc.) and unstructured data (imaging data). Given multimodal medical data $Data = \{I, C, Y\}$, where $I \in \mathbb{R}^N$ represents CT image data, $X \in \mathbb{R}^{N \times D_{image}}$ is the image feature matrix containing ND -dimensional feature vectors, $C \in \mathbb{R}^{N \times D_{clinical}}$ is the corresponding clinical data containing numerical values and categorical features, $Y \in \mathbb{R}^{N \times 2}$ is the label matrix indicating whether the *KRAS* gene is mutated, and N represents the total number of patients.

In this study, we consider the *KRAS* gene mutation prediction task as a node classification problem for heterogeneous graphs in a multi-view framework, where the diversity of edge types mainly represents heterogeneity. Formally, define the heterogeneous graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{R})$, where \mathcal{V}, \mathcal{E} represent the set of nodes and the set of edges, respectively, and \mathcal{R} represents the set of relation types. Each edge $e \in \mathcal{E}$ is associated with their mapping function $\varphi(e): \mathcal{E} \rightarrow \mathcal{R}$. Meanwhile, $A = \{A_1, A_2, \dots, A_{|\mathcal{R}|}\}$ represents the set of adjacency matrices corresponding to each relation. In the multi-view setting, we define the corresponding relational subgraph $\mathcal{G}_r = (\mathcal{V}_r, \mathcal{E}_r)$ for each view r . The feature matrix and adjacency matrix of each view are defined as $Z_r \in \mathbb{R}^{N \times D_r}$ and $A_r \in \mathbb{R}^{N \times N}$.

Heterogeneous graph node construction

Specifically, in this study, the patient is regarded as a node in the heterogeneous graph, fusing CT image data and clinical data. Inception-ResNet v2 is used as the backbone architecture (46) to extract its feature representation for the patient's image data. Eventually, the initial node

attribute $Z \in \mathbb{R}^{N \times D}$ is obtained by concatenating the image feature representation $X \in \mathbb{R}^{N \times D_{image}}$ and the clinical feature representation $C \in \mathbb{R}^{N \times D_{clinical}}$.

$$Z = [X, C] \quad [1]$$

Heterogeneous graph relationship construction

Clinical data contain many aspects of patient information, providing physicians with a multidimensional perspective about patients and helping them understand the conditions more comprehensively (47,48). Therefore, inspired by (49), when constructing the edges of the heterogeneous graph, we fully consider the similarities between patients regarding imaging and clinical data to reflect their complex associations. In order to fully explore this association, we first use K-prototype clustering (50) to partition the clinical data C and obtain $|\mathcal{R}|$ categories of non-overlapping feature combinations, which represent different multi-relational semantics, as shown in Figure 2, and are denoted as $T = \{T_1, T_2, \dots, T_{|\mathcal{R}|}\}$. Then, different views are built based on these relational semantics. In other words, according to different clinical feature types, we split heterogeneous graphs with multiple relational semantics into a set of relational subgraphs $\mathcal{G} = \{\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_R\}$ under multiple views, in order to achieve effective decoupling of multiple semantic relations. The advantage of doing so is that it can retain more complete information under each relationship semantics and exhibit stronger representational capabilities, ensuring that each view contains only a single semantic type of relationship. Specifically, given $\mathcal{G}_r = (\mathcal{V}_r, \mathcal{E}_r)$, its adjacency matrix A_r can be represented as:

$$A_r(i, j) = \varphi(e) = s(x_i, x_j) \sum_{m=1}^M d(c_{i,m}^r, c_{j,m}^r) \quad [2]$$

where $s(\cdot)$ represents the image feature similarity between two patients, $d(\cdot)$ shows the phenotypic distance between the clinical features of the patients, x_i represents the i th row in X , $c^{(r)}$ represents the clinical features in each view r , and M represents the number of clinical feature types contained under the relational semantics of each view. $s(\cdot)$ is defined as follows:

$$s(x_i, x_j) = \exp\left(-\frac{u^2(x_i, x_j)}{2\sigma^2}\right) \quad [3]$$

where σ is the kernel parameter and $u(\cdot)$ is the correlation distance function, and the second half is the Pearson correlation coefficient (51), defined as follows:

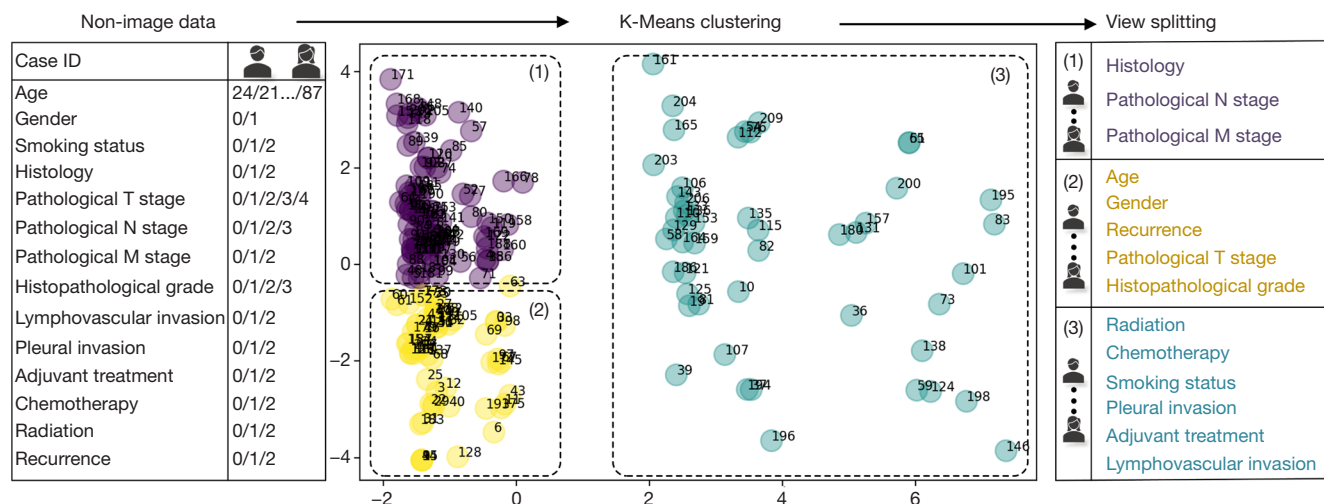


Figure 2 Decoupling of multiple relational semantics. T, tumor; N, node; M, metastasis.

$$u(x_i, x_j) = 1 - \frac{\text{cov}(x_i, x_j)}{\sigma(x_i)\sigma(x_j)} \quad [4]$$

where $\text{cov}(\cdot)$ is the covariance and $\sigma(\cdot)$ is the standard deviation.

The clinical features $C \in \mathbb{R}^{N \times D_{\text{clinical}}}$ contain both numerical and categorical features. For numerical type features, $d(\cdot)$ is defined as a unit step function:

$$d(c_{i,m}^r, c_{j,m'}^r) = \begin{cases} 1, & \text{if } |c_{i,m}^r - c_{j,m'}^r| < \theta \\ 0, & \text{otherwise.} \end{cases} \quad [5]$$

where θ is an adjustable parameter. For categorical features, define $d(\cdot)$ as the Kronecker function:

$$d(c_{i,m}^r, c_{j,m'}^r) = \begin{cases} 1, & \text{if } c_{i,m}^r = c_{j,m'}^r \\ 0, & \text{otherwise.} \end{cases} \quad [6]$$

Thereby, we construct \mathcal{G}_r ($r=1, 2, 3$) under the multi-view relationship semantics and use them as inputs to the subsequent network.

Single-view graph representation learning

After completing the heterogeneous graph construction, this section details the specific process of MASA-HGN for graph representation learning within a single view. First, we enrich the representation of patient nodes by enhancing the embedding of central nodes. In addition, we introduce a node embedding attention mechanism to consider the

differential contribution of different nodes' own features in updating the embedding. Then, for each relational type, we assign an aggregation module to perform aggregation operations in each relational subgraph using GCNs to generate neighborhood information specific to each view, which is crucial for the effectiveness and robustness of the model. *Figure 3* illustrates the procedure.

Central node embedding attention (CNE)

Node embedding enhancement aims to enhance the representation of a node by transforming the patient node's own features into a new feature space that captures the inherent information of the node. Node embedding is enhanced through the following process:

$$\widetilde{Z}_i^{c,l} = \sigma(I_i Z_i^{l-1} W^{c,l-1}) \quad [7]$$

where $\widetilde{Z}_i^{c,l}$ is the learned node embedding of the center node v_i at layer l , $W^{c,l-1}$ is the learnable parameter matrix for the feature transformation of the center node, and I is the unit matrix. In particular, at $l=1$, information aggregation is performed in conjunction with the initial attribute z_i^0 of the patient node.

At the same time, we consider the importance of the center node's own embedding in the updating process, so we need to calculate its contribution. According to $Z^l \in \mathbb{R}^{N \times D^l}$, let $\widetilde{Z}_i^{c,l}$ be the enhanced embedding of the i th patient node in Z^l , and calculate the weight of the own embedding of each patient node through the softmax function; the specific process is as follows:

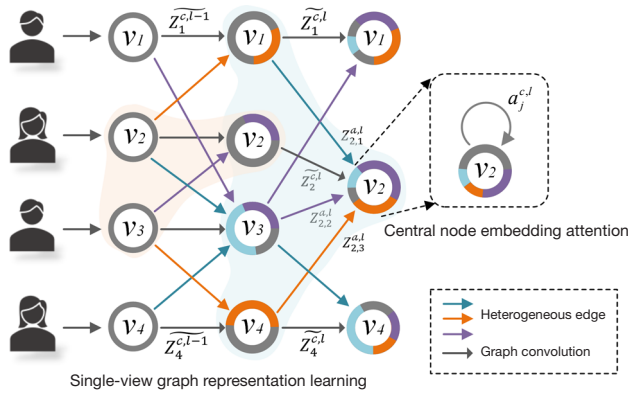


Figure 3 Decoupling of multiple relational semantics.

$$a_i^{c,l} = \text{softmax} \left(\lambda^T \cdot \tanh \left(W^{c,l} \widetilde{Z}_i^{c,l} + b^{c,l} \right) \right) \\ = \frac{\exp \left(\lambda^T \cdot \tanh \left(W^{c,l} \widetilde{Z}_i^{c,l} + b^{c,l} \right) \right)}{\sum_{n \in \mathcal{R}_i} \exp \left(\lambda^T \cdot \tanh \left(W^{c,l} \widetilde{Z}_i^{c,l} + b^{c,l} \right) \right)} \quad [8]$$

where $W^{c,l}$ and $b^{c,l}$ are the training weights and bias, respectively. $\tanh(\cdot)$ is used to adjust the network output to the range $(-1, 1)$ to prevent saturation during training. λ is a trainable weight vector.

Neighborhood node aggregation

The goal of node-level aggregation representation is to capture the neighborhood information of the target node. For the relational subgraph under each semantics, we capture the knowledge of the same edge semantics and generate adjacency embedding through information aggregation to reduce redundant information. In each view r , the following aggregation process is implemented based on the adjacency matrix A_r :

$$Z_{i,r}^{a,l} = \sigma \left(\widetilde{D}^{-\frac{1}{2}} \left(\widetilde{A}_r \right)_i \widetilde{D}^{-\frac{1}{2}} Z_r^{l-1} W_r^{a,l-1} \right) \quad [9]$$

where $Z_{i,r}^{a,l}$ is the neighborhood information embedding learned by node v_i in the l -layer r view. \widetilde{A}_r is the normalized adjacency matrix with the self-connections removed, D is the diagonal matrix, $\widetilde{D}_{ii} = \sum_j \widetilde{A}_{ij}$, $W_r^{a,l-1}$ is the adjacency weight matrix of the $l-1$ layer, and $\sigma(\cdot)$ is the activation function.

Multi-view heterogeneous information fusion

After completing the graph representation learning within

a single view, it is necessary to aggregate the neighborhood embeddings of different semantic relations to learn a comprehensive representation of the nodes. However, simply performing an average or summation operation on the neighborhood embeddings of all views may ignore the deep semantic information contained in the nodes and relations in heterogeneous graphs. Therefore, it becomes essential to integrate this semantic information effectively. To address this challenge, we propose a novel attention mechanism that comprehensively captures the heterogeneity information of specific semantic neighborhood structures and heterogeneous relation types. Specifically, we design two complementary attention mechanisms, node attribute semantic-aware attention (NASA) and heterogeneous relationship semantic-aware attention (HRSA). NASA focuses on the node's own feature performance under different views, whereas HRSA concentrates on the interaction effects among nodes due to different relationship types. This dual attention mechanism provides a new perspective for understanding complex inter-patient relationships, and an effective solution for multi-view information fusion in gene mutation prediction tasks. Figure 4 illustrates the specific flow of information fusion among multiple views.

NASA

Node semantics itself is differentially affected by neighboring nodes under different relationship semantics. To quantify this influence, we introduce the node semantics-aware attention mechanism. This mechanism can measure the impact of neighborhood embedding on the target node and is defined using contextual information about the neighborhood information:

$$a_{i,r}^{l,Se} = \frac{\exp \left(\text{LeakyReLU} \left(a_{Se}^l \left[\ell_2 \left(Z_{i,r}^{c,l} \right) \right] \ell_2 \left(Z_{i,r}^{a,l} \right) \right) \right)}{\sum_{n \in \mathcal{R}_i} \exp \left(\text{LeakyReLU} \left(a_{Se}^l \left[\ell_2 \left(Z_{i,r}^{c,l} \right) \right] \ell_2 \left(Z_{i,n}^{a,l} \right) \right) \right)} \quad [10]$$

where $a_{i,r}^{l,Se}$ represents the node semantic-aware attention score of node v_i in the l -layer r -view, a_{Se}^l is the vector of semantic-aware attention of the node that needs to be learned, and ℓ_2 stands for the L2 normalization for rescaling the distance between $Z_{i,r}^{c,l}$ and $Z_{i,n}^{a,l}$. \mathcal{R}_i stands for the relation set that is connected to the target node v_i . With NASA, we effectively capture the semantic dependencies retained in the nodes themselves.

HRSA

For different heterogeneous relationships, nodes are

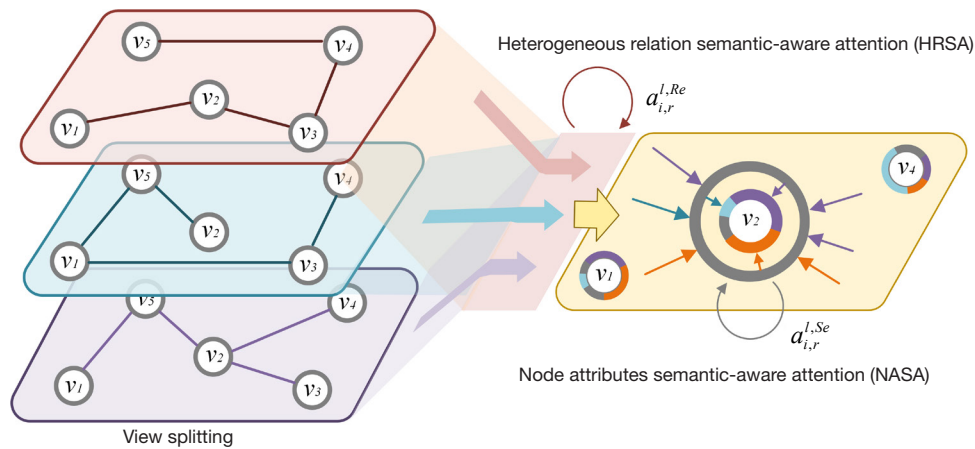


Figure 4 Specific process of heterogeneous information fusion between multiple views.

surrounded by nodes of multiple relationship types, and the contribution of each relationship to a node varies. Therefore, it is necessary to construct a metric that can change with nodes and relationship types. Unlike the node attribute semantics-aware attention mechanism, the heterogeneous relationship semantics-aware attention should be orthogonal to the context of the neighborhood embedding. To achieve this goal, we assign a learnable vector $k_r \in \mathbb{R}^{N \times D_r}$ to each relationship semantic $r \in \mathcal{R}$ at each layer, which is computed as follows:

$$a_{i,r}^{l,Re} = \frac{\exp(\text{LeakyReLU}(a_{Re}^l \cdot k_r^l))}{\sum_{n \in \mathcal{N}_i} \exp(\text{LeakyReLU}(a_{n,Re}^l \cdot k_r^l))} \quad [11]$$

where $a_{i,r}^{l,Re}$ represents the heterogeneous relation-aware attention score of node v_i in the l -layer r -view and a_{Re}^l represents the heterogeneous relation-aware attention vector. In this way, the association between each heterogeneous relation and neighborhood embedding is decoupled, and edge heterogeneous dependencies are effectively captured.

Inter-view adjacent embedding attention

After obtaining the NASA and HRSA under each view, they are effectively combined as each view-specific attention score. In this process, the rich semantic and heterogeneous relational information contained in each node in its local structure is deeply analyzed, aiming to comprehensively capture and utilize this information to more accurately depict the semantic features and relational dependencies of the nodes. To achieve this goal, we introduce a balance coefficient ω that dynamically adjusts and balances the weights of NASA and heterogeneous

relational semantic-aware attention in the learning process of node representations, so that the node attribute semantic information and heterogeneous relational semantic information can be appropriately fused. Finally, the specific calculation process of view weights for each patient is as follows:

$$a_{i,r}^{a,l} = \omega \cdot a_{i,r}^{l,Se} + (1 - \omega) \cdot a_{i,r}^{l,Re} \quad [12]$$

where the balance coefficient ω is a learnable parameter.

Multi-view fusion module

In this section, we illustrate the implementation details of the multi-view fusion module. The module introduces the CNE mechanism and the inter-view adjacent embedding attention mechanism. Based on the patient node v_i 's own embedding in different relational semantics, its semantic representation and heterogeneous information under multiple views are effectively aggregated, by which a richer and more comprehensive node representation is constructed, in addition, in order to facilitate the training and reduce the interference of the node degree, the L2 normalization technique is used. Finally, we obtain the node representation based on the fusion output of the attention mechanism:

$$Z_i^{l+1} = \ell_2 \left(\sigma \left(a_i^{c,l} \cdot Z_i^{c,l} + \sum_{r \in \mathcal{R}_i} a_{i,r}^{a,l} \cdot Z_{i,r}^{a,l} \right) \right) \quad [13]$$

Z_i^{l+1} is the final node representation learned at layer $l+1$, $Z_i^{c,l}$ is the CNE, and $Z_{i,r}^{a,l}$ is the adjacent information embedding. σ is the activation function and ℓ_2 denotes the normalization operation.

Algorithm 1 Training procedure of MVASA-HGN

Input: The multimodal heterogeneous graph $\mathcal{G}=(\mathcal{V},\mathcal{E},\mathcal{R})$

 The initial node feature $\{x_i, \forall i \in \mathcal{V}\}$

 The view set $\{\gamma_1, \gamma_2, \dots, \gamma_{|\mathcal{R}|}\}$

 The number of layers L

 The weight coefficient ω

Output: The final node embedding Z_i

1 for $r=1, \dots, \mathcal{R}$ do

2 for $l=1, \dots, L$ do

3 for $i \in \mathcal{V}$ do

4 Node-self hidden state $\widetilde{Z}_i^{c,l} \leftarrow \sigma(I_l Z_i^{l-1} W^{c,l-1})$

5 Calculate central node embedding attention $a_i^{c,l}$

6 Aggregate neighborhood hidden state

$$Z_{i,r}^{a,l} \leftarrow \sigma\left(\widetilde{D}^{-\frac{1}{2}}(\widetilde{A}_r)_i \widetilde{D}^{-\frac{1}{2}} Z_r^{l-1} W_r^{a,l-1}\right)$$

7 Calculate node attributes semantic-aware attention $a_{i,r}^{l,Se}$

8 Calculate heterogeneous relation semantic-aware attention $a_{i,r}^{l,Re}$

9 Calculate inter-view adjacent embedding attention score

$$a_{i,r}^{a,l} \leftarrow \omega \cdot a_{i,r}^{l,Se} + (1-\omega) \cdot a_{i,r}^{l,Re}$$

10 end

11 end

12 Fuse hidden state for all views at $l+1$ layer

$$Z_i^{l+1} \leftarrow \ell_2\left(\sigma\left(a_i^{c,l} \cdot Z_i^{c,l} + \sum_{r \in \mathcal{R}} a_{i,r}^{a,l} \cdot Z_{i,r}^{a,l}\right)\right)$$

13 end

14 Calculate cross-entropy loss

$$L_{class} \leftarrow -\sum_{i=1}^N \left(-y_i \log \widehat{y}_i - (1-y_i) \log (1-\widehat{y}_i)\right)$$

15 return $Z_i \leftarrow Z_i^L, \forall i \in \mathcal{V}$

Loss function

In Algorithm 1, the training process of the proposed MVASA-HGN model is elaborated. The multi-view fusion module can obtain the output embedding $Z \in \mathbb{R}^{N \times D^L}$. For the final node classification task, we introduce the cross-

entropy loss function. Denote $Z_i \in \mathbb{R}^{N \times D^L}$ to represent the node embedding of the i th patient in Z . A fully connected layer with a softmax activation function is used to obtain the predicted classification result \widehat{y}_i . The definition is as follows:

$$\widehat{y}_i = \text{softmax}(w_i^T Z_i + b_i) \quad [14]$$

where $w_i^T \in \mathbb{R}^{N \times D^L}$ and b_i are the weights and biases in the fully connected layer. For the training process containing N patients, the loss function is defined as follows:

$$L_{class} = -\sum_{i=1}^N \left(-y_i \log \widehat{y}_i - (1-y_i) \log (1-\widehat{y}_i)\right) \quad [15]$$

y_i is the label of the KRAS gene mutation of the i th patient.

Results**Data acquisition**

This study used data from 363 patients with information on *KRAS* gene mutations collected from cooperative hospitals in Shanxi Province as training and validation sets. Testing data was obtained from the public dataset NSCLC Radiogenomics (52) available on The Cancer Imaging Archive (TCIA) website (<https://wiki.cancerimagingarchive.net>), which contains imaging data, genetic data, and clinically informative data from 211 patients with NSCLC. After screening, a total of 168 cases were included to construct the testing dataset, and the demographic and clinical information of these patients is shown in *Table 1*, where data in parentheses are percentages unless otherwise stated. Patients were further categorized into *KRAS* mutant (1, mutant) and *KRAS* wild type (0, wild) based on clinical information, and some examples are shown in *Figure 5*.

Tumor regions of interest in CT images were manually outlined by a radiologist with 10 years of experience in chest imaging at the cooperative hospital and reviewed by a radiologist with 15 years of experience. In case of disagreement, the final result was determined by the agreement of the two radiologists.

Implementation details

The experiments in this study were all implemented using the Pytorch framework on a workstation equipped with NVIDIA RTX A6000 GPUs (NVIDIA, Santa Clara, CA, USA). The Adam optimizer was used to adjust the learning rate with the initial learning rate set to 0.0001, dropout rate to 0.5, and epoch to 400. θ was set to 2.

Table 1 Clinical characteristics of patients

Patient characteristics	Hospital (n=363)	TCIA (n=168)
Histology		
LUAD	257 (70.8)	151 (89.9)
LUSC	106 (29.2)	17 (10.1)
Median age [range] (years)	62 [43–80]	68 [24–87]
Gender		
Male	156 (43.0)	106 (63.1)
Female	207 (57.0)	62 (36.9)
Smoking status		
Never	127 (35.0)	42 (25.0)
Former/light	189 (52.1)	101 (60.1)
Current/heavy	47 (12.9)	25 (14.9)
T stage		
T1	138 (38.0)	58 (34.5)
T2	119 (32.8)	44 (26.2)
T3	71 (19.6)	15 (8.9)
T4	35 (9.6)	4 (2.4)
Not collected	0	47 (28.0)
N stage		
N0	276 (76.0)	95 (56.5)
N1	64 (17.7)	12 (7.1)
N2	23 (6.3)	14 (8.3)
Not collected	0	47 (28.0)
M stage		
M0	302 (83.2)	116 (69.0)
M1	61 (16.8)	5 (3.0)
Not collected	0	47 (28.0)
Histopathological grade		
G1	87 (24.0)	27 (16.0)
G2	201 (55.4)	71 (42.3)
G3	75 (20.6)	23 (13.7)
Not collected	0	47 (28.0)
Lymphovascular invasion		
Absent	311 (85.7)	100 (59.5)
Present	52 (14.3)	16 (9.5)
Not collected	0	52 (31.0)

Table 1 (continued)**Table 1** (continued)

Patient characteristics	Hospital (n=363)	TCIA (n=168)
Pleural invasion		
Yes	90 (24.8)	35 (20.8)
No	273 (75.2)	86 (51.2)
Not collected	0	47 (28.0)
Adjuvant treatment/chemotherapy		
Yes	48 (13.2)	42 (25.0)
No	315 (86.8)	125 (74.4)
Not collected	0	1 (0.6)
Radiation		
Yes	66 (18.2)	15 (8.9)
No	297 (81.8)	152 (90.5)
Not collected	0	1 (0.6)
Recurrence		
Yes	146 (40.2)	44 (26.2)
No	217 (59.8)	123 (73.2)
Not collected	0	1 (0.6)

Data are presented as n (%) unless otherwise noted. LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; T, tumor; N, node; M, metastasis; TCIA, The Cancer Imaging Archive.

In order to efficiently process the clinical data, the following strategy was adopted. For numerical features in structured data, the features were scaled to the range of 0 to 1 using the Z-score normalization method. One-hot coding was used to convert classification features into binary vectors for subsequent data analysis. For the CT image data, in order to eliminate the ray differences between different CT scanners, 168 case images were resampled to obtain 1×1×1 mm³ isotropic voxels, and a 64×64×64 VOI region was cropped to obtain a 64×64×64 VOI region based on the labeled lesion area.

In addition, to assess the performance of the model, we calculated accuracy (ACC), sensitivity (SEN), specificity (SPE), precision (PRE), and F1 scores, as well as the area under the receiver operating characteristic curves (AUCs) of the cases, which are defined as:

$$SEN = \frac{TP}{TP + FN} \quad [16]$$

$$SPE = \frac{TN}{TN + FP} \quad [17]$$

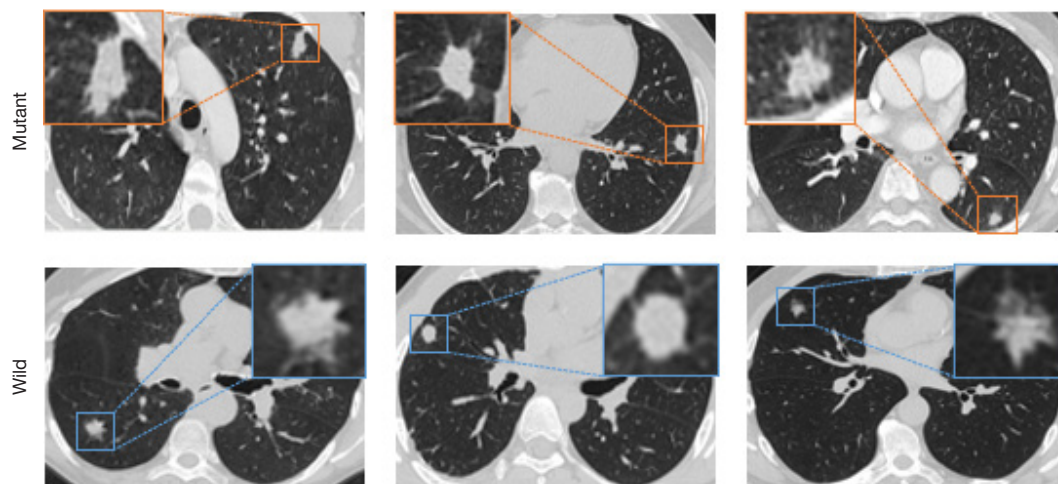


Figure 5 Examples of KRAS mutant and KRAS wild-type samples. Orange boxes represent KRAS mutant tumor regions and blue boxes represent KRAS wild-type tumor regions.

$$ACC = \frac{TP + TN}{TP + FP + TN + PN} \quad [18]$$

$$PRE = \frac{TP}{TP + FP} \quad [19]$$

$$F1\ score = \frac{2 \times PRE \times SEN}{PRE + SEN} \quad [20]$$

where TP, TN, FP, and FN stand for true positive, true negative, false positive, and false negative, respectively.

Ablation experiment

Effectiveness of multiple views

In this paper, we decompose the heterogeneous graph into multiple views from the perspective of clinical information types and learn the node representations in each relational subgraph, ultimately integrating the synthesized representations from all the views to predict the mutation status of patients' *KRAS* genes. To validate the effectiveness of the joint multi-view heterogeneous relationship prediction framework, the following network components are implemented:

- (I) Three single-view models are developed based on imaging and clinical features, each focusing on only one relationship type. Each model is classified using a three-layer graph convolutional layer and a fully connected layer with an activation function.
- (II) Combining two single-view models with information fusion using attention, each view is classified using three graph convolution layers and

fully connected layers with activation functions.

- (III) MVASA-HGN, proposed in this paper, adaptively aggregates heterogeneous information from the three views for node classification.

As shown in *Table 2*, the experimental results indicate that combining all types of clinical features to construct an MVASA-HGN achieves the best results compared to single-view and combined two-view heterogeneous graph networks. We also observed that the performance of the view 3 alone, as well as the classification performance when combining view 3 with other views, is superior to other network components. This indirectly suggests that the clinical features in view 3 are instructive for the prediction of the *KRAS* gene mutation status. Our method significantly outperforms other network components in terms of ACC, SEN, SPE, PRE, and F1 scores, which fully demonstrates the validity of the proposed multi-view heterogeneous relationship.

Effectiveness of attention mechanisms

In Sections “Ablation experiment” and “Hyperparameter sensitivity”, we presented three attention mechanisms: CNE, NASA, and HRSA. CNE evaluates the importance of a node's own features. Meanwhile, NASA and HRSA collaboratively capture the heterogeneity information preserved in the graph structure in terms of both node and relation semantics, effectively weighing the contribution of adjacent embeddings across different views. To assess the effectiveness of these attention components, we devised the following experimental setup:

- (I) Removing CNE, where the output node features

Table 2 Results of ablation experiments for multi-view effectiveness

Methods	SEN (%)	SPE (%)	ACC (%)	PRE (%)	F1 score (%)	AUC (%)
GCN [1]	78.23	80.18	75.55	79.77	79.00	78.43
GCN [2]	74.51	78.39	71.64	76.43	75.46	72.87
GCN [3]	82.12	85.66	82.34	83.89	83.00	85.16
MVASA-HGN						
[1] & [2]	81.77	85.45	81.06	83.88	82.82	84.40
[1] & [3]	84.11	87.98	84.54	87.57	85.81	88.79
[2] & [3]	83.49	86.32	83.99	85.62	84.55	86.55
[1] & [2] & [3] (ours)	85.71	89.67	85.29	88.24	86.96	91.94

GCN, graph convolutional network; MVASA-HGN, multi-view adaptive semantics-aware heterogeneous graph network; SEN, sensitivity; SPE, specificity; ACC, accuracy; PRE, precision; AUC, area under the curve.

Table 3 Results of ablation experiments on the effectiveness of attention mechanisms

Methods	SEN (%)	SPE (%)	ACC (%)	PRE (%)	F1 score (%)	AUC (%)
CNE	79.50	83.69	78.96	83.16	81.31	84.33
NASA	83.89	89.84	83.67	86.99	85.42	89.85
HRSA	81.58	87.31	80.90	85.28	83.41	86.99
Ours	85.71	89.67	85.29	88.24	86.96	91.94

CNE, central node embedding attention; NASA, node attribute semantic-aware attention; HRSA, heterogeneous relation semantic-aware attention; SEN, sensitivity; SPE, specificity; ACC, accuracy; PRE, precision; AUC, area under the curve.

from the last layer are directly and adaptively integrated with aggregated adjacent embedding from various views as the updated node representation; other parts remain unchanged.

- (II) Removing NASA, where the adjacent embeddings from different views are aggregated using only HRSA; other parts remain unchanged.
- (III) Removing HRSA, where the adjacent embeddings of different views are aggregated using only NASA; other parts remain unchanged.

The experimental results are shown in *Table 3*. We can see that each of the designed attention components positively affect the model performance, and the model achieves optimal classification accuracy when adaptively combining adjacent embedding and the node's own embedding under different views through the three attention mechanisms. Notably, we observed that node embedding attention significantly impacted model performance, demonstrating the importance of effectively combining the node's own features with neighbor features in constructing node representations. In addition, compared to using only

NASA, using only heterogeneous relational semantic-aware attention had better performance, suggesting that heterogeneous relational semantic is more effective than node semantic in identifying discriminative features. However, MVASA-HGN achieved the best performance when two complementary attentions collaborated.

Hyperparameter sensitivity

In this section, the effects of three hyperparameters in the model, which are the number of clusters \mathcal{R} , the balance coefficient ω , and the number of graph convolution layers l , were mainly analyzed to demonstrate the effectiveness and robustness of the proposed MVASA-HGN.

Cluster number \mathcal{R}

The clustering number \mathcal{R} represents the number of relationship types in the heterogeneous graph and directly determines the complexity of inter-patient associations. To obtain the optimal number of relationship types, we varied the value of \mathcal{R} and observed the model's performance.

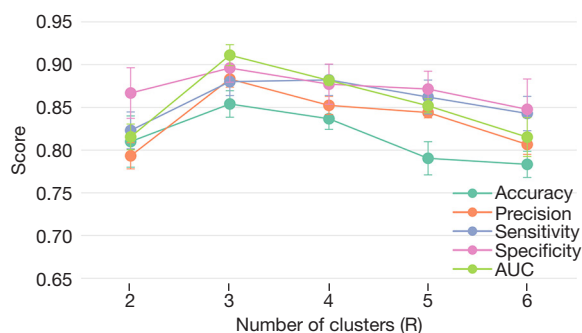


Figure 6 Effect of clustering number R . AUC, area under the curve.

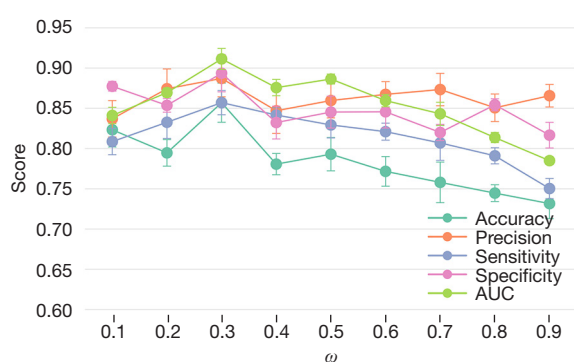


Figure 7 Effect of balance coefficient ω . AUC, area under the curve.

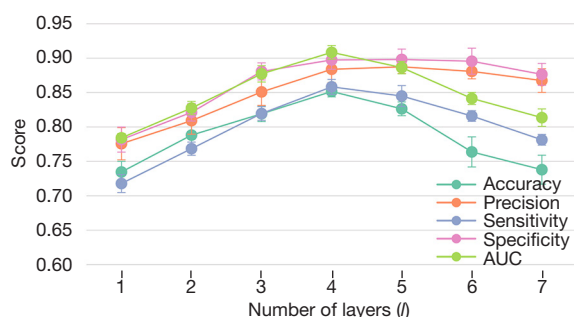


Figure 8 Effect of the number of graph convolution layers l . AUC, area under the curve.

Figure 6 shows the performance metrics with different R . The experimental results indicated that the optimal performance is achieved when $|\mathcal{R}|=3$. When \mathcal{R} is too small or too large, the model's performance degrades. This is because the model cannot model complex relationships when \mathcal{R} is too small. Meanwhile, if \mathcal{R} is too large, it results in redundancy of information between multiple relations, bringing meaningless noise to the network and leading to

performance degradation. Therefore, in this paper, $|\mathcal{R}|=3$ is adopted as the number of relationship types.

The balance coefficient ω

The balance coefficient quantifies the relative contribution of NASA and HRSA in the model. As seen in Figure 7, the model achieved the best performance at $\omega=0.3$. It is worth noting that a larger ω represents a higher weight of the node's semantic-aware attention. Correspondingly, the weight of heterogeneous semantic-aware attention is lower. The optimal value of ω reveals from another perspective that heterogeneous relationship semantic may be more discriminative for predicting the *KRAS* gene mutation status.

Number of graph convolution layers l

To evaluate the impact of the number of graph convolution layers on model performance, we tested the performance of the model by changing the number of layers. The experimental results are shown in Figure 8. The results indicate that as the number of layers increases, the model can capture more information, leading to a gradual enhancement in performance. However, as the number of layers increases further, the performance fluctuates and tends to decline, possibly due to over-smoothing caused by an excess of layers. Therefore, in this paper, $l=4$ is chosen as the number of layers for graph convolution.

Comparison experiments

Existing mutation prediction methods

In this section, we compared our model to three recently published DL methods (53-55) for predicting gene mutation status in lung cancer. The effectiveness of our proposed method was evaluated, and the experimental results are shown in Table 4. The results show that our method outperforms these methods, achieving at least 3.71% higher ACC, 6.77% higher SPE, and 5.69% higher SEN. This suggests that our method is highly effective in accurately determining the *KRAS* gene mutation status. It is worth noting that Xiong *et al.* (54) improved the AUC from 77.6% to 83.8% after adding the clinical information, which further supports our point that the inclusion of clinical information is crucial for enhancing the predictive performance of the model.

Comparison with existing GNN-based approaches

To further assess the classification performance of the models, in this section, we compared the performance of two

existing heterogeneous GNN classification methods (56,57) and three multi-view GCN classification methods (58-60) in predicting the mutation status of *KRAS* genes for NSCLC patients. The specific methods are described below:

- (I) Heterogeneous graph transformer (HGT) (56): an HGT architecture designed with node and edge type-related parameters to characterize the heterogeneous attention of each edge.
- (II) SimpleHGN (57): a heterogeneous GNN built on the GAT backbone, improving the HGNN with learnable relational embeddings and residual connections.
- (III) SpectralGCN (58): in this study, it is extended to three branches, each comprising a SpectralGCN with three graph convolutional layers. The outputs from these branches are fused through an average pooling layer and classified using a fully connected layer.
- (IV) Co-GCN (59): a multi-view GCN model capable of adaptively extracts graph information from multiple views by employing the combined Laplace

operator.

- (V) MSGFN (60): a multi-stage graph fusion GCN model. In this study, MSGFN is used to realize the classification task by fusing information from relational subgraphs under multiple views.

The experimental results are shown in *Table 5*. The results show that compared with the multi-view methods (56-58), our method shows significant advantages in all the metrics, which fully verified the rationality and effectiveness of our constructed view and multi-view information fusion strategy. Similarly, compared with heterogeneous graph methods (56,57), our method also achieves good performance, showing that the hierarchical processing strategy combining the multi-view idea can effectively utilize these complex heterogeneous relationships. In summary, our MVASA-HGN method outperforms existing GNN-based methods and achieves good performance in ACC, SEN, SPE, PRE, and F1 score.

Discussion

Rationalization of the graphical representation

In clinical medication decision-making, CT images and patients' baseline clinical characteristics form a crucial basis for physicians' decisions (48,61). These two types of heterogeneous information reveal the patient's health status from different and complementary perspectives, with CT images providing a full range of information about the lesion area. At the same time, the clinical characteristics reflect the patient's overall condition. Notably, inter-patient correlations also provide valuable clues for patient stratification (62,63), and existing studies have shown that

Table 4 Performance comparison with existing mutation prediction methods

Methods	SEN (%)	SPE (%)	ACC (%)	AUC (%)
Wang <i>et al.</i> (53)	72.27	75.41	73.86	81.00
Xiong <i>et al.</i> (54)	75.80	79.10	77.20	83.80
Jia <i>et al.</i> (55)	80.02	82.90	81.58	88.61
Ours	85.71	89.67	85.29	91.94

SEN, sensitivity; SPE, specificity; ACC, accuracy; AUC, area under the curve.

Table 5 Performance comparison with existing GNN-based methods

Methods	SEN (%)	SPE (%)	ACC (%)	PRE (%)	F1 score (%)	AUC (%)
HGT (56)	78.95	80.33	78.92	79.13	78.18	83.69
SimpleHGN (57)	83.73	87.18	83.05	86.89	85.28	89.17
SpectralGCN (58)	74.85	80.07	76.56	78.71	77.74	78.96
Co-GCN (59)	75.89	82.29	74.08	80.62	78.22	81.41
MSGFN (60)	80.33	84.65	82.11	84.05	82.16	88.23
Ours	85.71	89.67	85.29	88.24	86.96	91.94

GNN, graph neural network; HGT, heterogeneous graph transformer network; SimpleHGN, improved heterogeneous graphs based on graph attention networks; SpectralGCN, graph neural networks processing data by convolutional operations in the spectral domain; Co-GCN, co-training graph convolutional network; MSGFN, a multi-stage graph fusion graph convolutional network; SEN, sensitivity; SPE, specificity; ACC, accuracy; PRE, precision; AUC, area under the curve.

Table 6 Validity of heterogeneous graph construction

Node construction/edge construction	SEN (%)	SPE (%)	ACC (%)	PRE (%)	F1 score (%)	AUC (%)
Image feature/image feature similarity	61.98	72.31	61.58	73.25	67.15	69.33
Image feature similarity	81.82	84.49	82.99	84.45	83.11	87.79
Clinical features	83.74	86.66	83.16	84.67	84.20	89.08
Ours	85.71	89.67	85.29	88.24	86.96	91.94

SEN, sensitivity; SPE, specificity; ACC, accuracy; PRE, precision; AUC, area under the curve.

patients with similar clinical characteristics tend to show similarity in treatment response (49,64). Especially when the sample size is limited, fully considering such inter-patient correlations can significantly improve the predictive performance of the model.

KRAS mutation is one of the most frequently reported functionally acquired oncogenic driver mutations in NSCLC patients, present in 25–30% of lung adenocarcinoma cases (65,66). According to the recommendations of the National Comprehensive Cancer Network (NCCN) guidelines (67), pre-treatment detection of oncogenic driver mutations in patients with advanced NSCLC can more effectively guide the development of treatment strategies. Therefore, in this paper, we formulated a prediction of the *KRAS* mutation status of NSCLC patients as a heterogeneous graph node classification problem, and innovatively proposed a graph structure to fuse multimodal data to characterize the relationships among patients naturally. Specifically, patients' imaging and clinical features were taken as node attributes, whereas similarities between patients in imaging and clinical features form the edges of the graph network, and the complex heterogeneous relationships were hierarchically handled by constructing multiple views. Within each view, we further utilized the GCN to construct adjacent embedding of the same relationship types and augmented embedding of the nodes. Then, by adaptively fusing information from multiple views, the network can focus on both node semantics and heterogeneous relationship semantics to generate a comprehensive representation of the nodes. This approach effectively combines the patient's image and clinical data, and deeply explores the potential relationships between patients, providing robust support for clinical decision-making.

Validity of heterogeneous graph construction

In this section, we investigated the effectiveness of heterogeneous graph construction, mainly analyzing two

aspects: node construction and edge construction. Regarding node features, we fused the multimodal features of patient images and clinical information as the node features. In order to verify the rationality of this node feature construction, ablation experiments were performed on these two features. In terms of edge construction, we comprehensively considered two dimensions, similarity of image features and phenotypic distance between clinical features, to assess the similarity between nodes. The rationality and robustness of the edge construction were also demonstrated through the ablation experiments to ensure the reliability and validity of the whole heterogeneous graph construction process. The experimental results are shown in *Table 6*.

Based on the experimental results, we observed that combining both imaging and clinical features plays a crucial role in both the node and edge construction process, and this cross-modal information fusion strategy not only significantly improves the performance of the model, but also effectively enhances the accurate description of the complex heterogeneity among patient nodes by complementing each other. This construction method effectively improves the quality of the heterogeneous graph and provides a more accurate and solid foundation for subsequent data analysis and prediction.

Interpretability

Since the attention mechanism in MVASA-HGN captures the importance of each relational semantics for *KRAS* gene mutation prediction, it enables MVASA-HGN to explain which combination of features represents the type of relationship that plays a key role in decision-making. We calculated the average attention scores for all patients under each view during training and plotted the changes in attention scores in *Figure 9*. As the training progressed, the attention values of the three views were changing. Eventually, the attention weights were concentrated in view 3, which indicated the leading role of radiation,

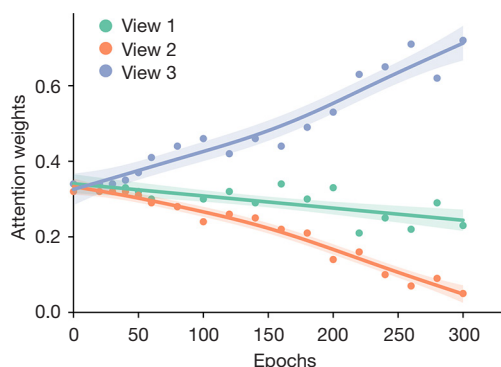


Figure 9 Importance of different relationship semantics. Three views represent relational subgraphs under different heterogeneous relationships.

chemotherapy, smoking status, pleural invasion, adjuvant treatment, and lymphovascular invasion information in the *KRAS* mutation status prediction. Comparatively, the attentional weights of the view 1 and views 2 decreased as the training phase progressed. This result not only demonstrates the effectiveness of the attentional mechanism in the MVASA-HGN model, but also further confirms that the mechanism can adaptively adjust the allocation of attention among different views during the training process to optimize the overall performance of the model.

Visualization

In order to visually evaluate the performance of different methods, we used t-distributed Stochastic Neighbor Embedding (t-SNE) (68) to project the node embedding vectors into a two-dimensional space and color them according to the type of *KRAS* gene mutation. We compared the proposed method with the existing GNN-based approaches; the results are shown in Figure 10. There was no clear distinction between the two types using SpectralGCN, with a high degree of overlap between the two categories. Co-GCN also showed a high category overlap, presenting fuzzy boundaries between the categories, which makes it challenging to distinguish *KRAS* mutation status. Compared to these two methods, HGT, MSGFN, and SimpleHGN improved, with SimpleHGN performing better under the guidance of learnable relationships. However, SimpleHGN had closer node distances between the two categories, suggesting that it had lost heterogeneous information to some extent. In contrast, our proposed

MVASA-HGN presents a significant separation between the two categories and can successfully classify nodes into different categories with almost no overlapping data points, which fully validates the effectiveness of our model.

Clinical applications

In this study, we used a graph structure-based multimodal information fusion approach to model the decision-making process of physicians in clinical practice, for similar cases, the similarity (relationship) with previous cases is naturally taken into account during decision-making. The model presented in this paper offers potential clinical applications:

- ❖ Non-invasive prediction of *KRAS* gene mutation status: provides a non-invasive means of predicting *KRAS* gene mutation status from easily accessible CT images, and can be reused as a validation tool in clinical practice. This property is significant for clinical medication guidance and is expected to replace the traditional invasive biopsy and cytology.
- ❖ Real-time update and prediction: In actual clinical practice, when a new test sample appears, the imaging and clinical features of the new sample are first extracted and integrated into the trained graph structure and the adjacency matrix A_r and feature matrix Z_r in $\mathcal{G}_r = (r=1, \dots, \mathcal{R})$ can be updated in real-time, and the trained MVASA-HGN model will give the prediction results of the new sample, providing timely and effective support for clinical decision-making.

Conclusions

In this paper, a new MVASA-HGN is introduced for fusing multimodal medical data that allows deep modeling of the complex associations among NSCLC patients. Unlike approaches relying on predefined meta-paths, MVASA-HGN fully explores the rich semantic information preserved in the nodes and heterogeneous relations through a designed attention mechanism to effectively handle the heterogeneity of edges. This approach not only provides a new perspective for the fusion of image and non-image information, but also creates a new way to explore the potential connection between images and genes. More importantly, it provides a non-invasive and cost-effective solution for identifying *KRAS* mutation status, which

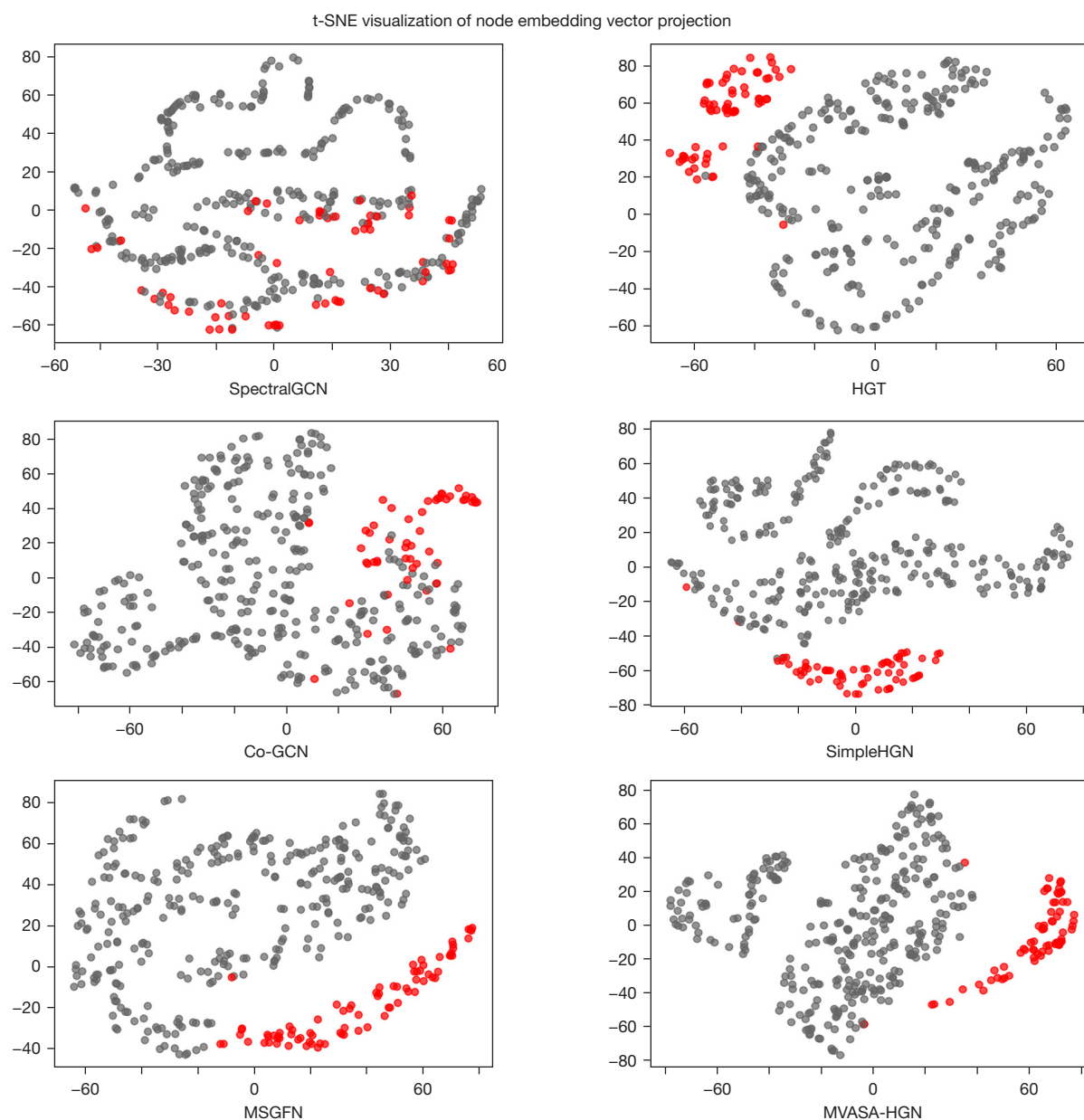


Figure 10 Visualization of different GNN methods. Red dots represent patients with *KRAS* gene mutations, and grey dots represent patients without *KRAS* gene mutations. SpectralGCN, graph neural networks processing data by convolutional operations in the spectral domain; HGT, heterogeneous graph transformer network; Co-GCN, co-training graph convolutional network; SimpleHGN, improved heterogeneous graphs based on graph attention networks; MSGFN, a multi-stage graph fusion graph convolutional network; MVASA-HGN, multi-view adaptive semantics-aware heterogeneous graph network; GNN, graph neural network.

provides strong support for subsequent treatment.

Acknowledgments

None.

Footnote

Funding: This work was supported by the National Natural Science Foundation of China (grant Nos. 61972274 and U21A20469), Applied Basic Research Project of Shanxi

Province, China (grant No. 202103021224066), and China Scholarship Council (grant No. 202306930022).

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://qims.amegroups.com/article/view/10.21037/qims-24-1370/coif>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by Shanxi Provincial People's Hospital Ethics Committee (No. 2023.299) and the requirement for individual consent for this retrospective analysis was waived.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Liu SM, Zheng MM, Pan Y, Liu SY, Li Y, Wu YL. Emerging evidence and treatment paradigm of non-small cell lung cancer. *J Hematol Oncol* 2023;16:40.
2. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. *CA Cancer J Clin* 2024;74:12-49. Erratum in: *CA Cancer J Clin* 2024;74:203.
3. Lu T, Yang X, Huang Y, Zhao M, Li M, Ma K, Yin J, Zhan C, Wang Q. Trends in the incidence, treatment, and survival of patients with lung cancer in the last four decades. *Cancer Manag Res* 2019;11:943-53.
4. Li Q, Zhou Q, Zhao S, Wu P, Shi P, Zeng J, Xiong X, Chen H, Kittaneh M, Bravaccini S, Zanoni M, Zhou C, Zhang J. KRAS mutation predict response and outcome in advanced non-small cell lung carcinoma without driver alterations receiving PD-1 blockade immunotherapy combined with platinum-based chemotherapy: a retrospective cohort study from China. *Transl Lung Cancer Res* 2022;11:2136-47.
5. Adderley H, Blackhall FH, Lindsay CR. KRAS-mutant non-small cell lung cancer: Converging small molecules and immune checkpoint inhibition. *EBioMedicine* 2019;41:711-6.
6. Nagasaka M, Li Y, Sukari A, Ou SI, Al-Hallak MN, Azmi AS. KRAS G12C Game of Thrones, which direct KRAS inhibitor will claim the iron throne? *Cancer Treat Rev* 2020;84:101974.
7. Ferrer I, Zugazagoitia J, Herbertz S, John W, Paz-Ares L, Schmid-Bindert G. KRAS-Mutant non-small cell lung cancer: From biology to therapy. *Lung Cancer* 2018;124:53-64.
8. Tang D, Kroemer G, Kang R. Oncogenic KRAS blockade therapy: renewed enthusiasm and persistent challenges. *Mol Cancer* 2021;20:128.
9. Abu Rous F, Singhi EK, Sridhar A, Faisal MS, Desai A. Lung Cancer Treatment Advances in 2022. *Cancer Invest* 2023;41:12-24.
10. Salem ME, El-Refai SM, Sha W, Puccini A, Grothey A, George TJ, Hwang JJ, O'Neil B, Barrett AS, Kadakia KC, Musselwhite LW, Raghavan D, Van Cutsem E, Tabernero J, Tie J. Landscape of KRAS(G12C), Associated Genomic Alterations, and Interrelation With Immuno-Oncology Biomarkers in KRAS-Mutated Cancers. *JCO Precis Oncol* 2022;6:e2100245.
11. Toki MI, Mani N, Smithy JW, Liu Y, Altan M, Wasserman B, Tuktamyshov R, Schalper K, Syrigos KN, Rimm DL. Immune Marker Profiling and Programmed Death Ligand 1 Expression Across NSCLC Mutations. *J Thorac Oncol* 2018;13:1884-96.
12. Liu C, Zheng S, Jin R, Wang X, Wang F, Zang R, Xu H, Lu Z, Huang J, Lei Y, Mao S, Wang Y, Feng X, Sun N, Wang Y, He J. The superior efficacy of anti-PD-1/PD-L1 immunotherapy in KRAS-mutant non-small cell lung cancer that correlates with an inflammatory phenotype and increased immunogenicity. *Cancer Lett* 2020;470:95-105.
13. Mok TSK, Lopes G, Cho BC, Kowalski DM, Kasahara K, Wu YL, de Castro G Jr, Turna HZ, Cristescu R, Aurora-Garg D, Loboda A, Luncford J, Kobie J, Ayers M, Pietanza MC, Piperdi B, Herbst RS. Associations of tissue tumor mutational burden and mutational status with clinical outcomes in KEYNOTE-042: pembrolizumab versus chemotherapy for advanced PD-L1-positive NSCLC. *Ann Oncol* 2023;34:377-88.
14. Boiarsky D, Lydon CA, Chambers ES, Sholl LM, Nishino M, Skoulidis F, Heymach JV, Luo J, Awad MM, Janne PA, Van Allen EM, Barbie DA, Vokes NI. Molecular

- markers of metastatic disease in KRAS-mutant lung adenocarcinoma. *Ann Oncol* 2023;34:589-604.
15. Rios Velazquez E, Parmar C, Liu Y, Coroller TP, Cruz G, Stringfield O, Ye Z, Makrigiorgos M, Fennessy F, Mak RH, Gillies R, Quackenbush J, Aerts HJWL. Somatic Mutations Drive Distinct Imaging Phenotypes in Lung Cancer. *Cancer Res* 2017;77:3922-30.
 16. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* 2016;278:563-77.
 17. Aonpong P, Iwamoto Y, Wang W, Lin L, Chen YW. Hand-crafted and deep learning-based radiomics models for recurrence prediction of non-small cells lung cancers. *Innovation in Medicine and Healthcare: Proceedings of 8th KES-InMed* 2020;1:135-44.
 18. Bera K, Braman N, Gupta A, Velcheti V, Madabhushi A. Predicting cancer outcomes with radiomics and artificial intelligence in radiology. *Nat Rev Clin Oncol* 2022;19:132-46.
 19. Liu Z, Duan T, Zhang Y, Weng S, Xu H, Ren Y, Zhang Z, Han X. Radiogenomics: a key component of precision cancer medicine. *Br J Cancer* 2023;129:741-53.
 20. Zhou M, Leung A, Echegaray S, Gentles A, Shrager JB, Jensen KC, Berry GJ, Plevritis SK, Rubin DL, Napel S, Gevaert O. Non-Small Cell Lung Cancer Radiogenomics Map Identifies Relationships between Molecular and Imaging Phenotypes with Prognostic Implications. *Radiology* 2018;286:307-15.
 21. Nair JKR, Saeed UA, McDougall CC, Sabri A, Kovacina B, Raidu BVS, Khokhar RA, Probst S, Hirsh V, Chankowsky J, Van Kempen LC, Taylor J. Radiogenomic Models Using Machine Learning Techniques to Predict EGFR Mutations in Non-Small Cell Lung Cancer. *Can Assoc Radiol J* 2021;72:109-19.
 22. Wang L, Wang S, Yu H, Zhu Y, Li W, Tian J. A Quarter-split Domain-adaptive Network for EGFR Gene Mutation Prediction in Lung Cancer by Standardizing Heterogeneous CT image. *Annu Int Conf IEEE Eng Med Biol Soc* 2021;2021:3646-9.
 23. Luo Y, Li S, Ma H, Zhang W, Liu B, Xie C, Li Q. CT-based decision tree model for predicting EGFR mutation status in synchronous multiple primary lung cancers. *J Thorac Dis* 2023;15:1196-209.
 24. Jia TY, Xiong JF, Li XY, Yu W, Xu ZY, Cai XW, Ma JC, Ren YC, Larsson R, Zhang J, Zhao J, Fu XL. Identifying EGFR mutations in lung adenocarcinoma by noninvasive imaging using radiomics features and random forest modeling. *Eur Radiol* 2019;29:4742-50.
 25. Morgado J, Pereira T, Silva F, Freitas C, Negrão E, de Lima BF, da Silva MC, Madureira AJ, Ramos I, Hespanhol V, Costa JL, Cunha A, Oliveira HP. Machine learning and feature selection methods for EGFR mutation status prediction in lung cancer. *Appl Sci* 2021;11:3273.
 26. Moreno S, Bonfante M, Zurek E, Cherezov D, Goldgof D, Hall L, Schabath M. A Radiogenomics Ensemble to Predict EGFR and KRAS Mutations in NSCLC. *Tomography* 2021;7:154-68.
 27. Sun L, Dong Y, Xu S, Feng X, Fan X. Predicting Multi-Gene Mutation Based on Lung Cancer CT Images and Mut-SeResNet. *Appl Sci* 2023;13:1921.
 28. Zhang H, He M, Wan R, Zhu L, Chu X. Establishment and Evaluation of EGFR Mutation Prediction Model Based on Tumor Markers and CT Features in NSCLC. *J Healthc Eng* 2022;2022:8089750.
 29. Yang X, Liu M, Ren Y, Chen H, Yu P, Wang S, Zhang R, Dai H, Wang C. Using contrast-enhanced CT and non-contrast-enhanced CT to predict EGFR mutation status in NSCLC patients-a radiomics nomogram analysis. *Eur Radiol* 2022;32:2693-703.
 30. Gao J, Niu R, Shi Y, Shao X, Jiang Z, Ge X, Wang Y, Shao X. The predictive value of [18F]FDG PET/CT radiomics combined with clinical features for EGFR mutation status in different clinical staging of lung adenocarcinoma. *EJNMMI Res* 2023;13:26.
 31. Kim S, Lim JH, Kim CH, Roh J, You S, Choi JS, Lim JH, Kim L, Chang JW, Park D, Lee MW, Kim S, Heo J. Deep learning-radiomics integrated noninvasive detection of epidermal growth factor receptor mutations in non-small cell lung cancer patients. *Sci Rep* 2024;14:922.
 32. Chen S, Han X, Tian G, Cao Y, Zheng X, Li X, Li Y. Using stacked deep learning models based on PET/CT images and clinical data to predict EGFR mutations in lung cancer. *Front Med (Lausanne)* 2022;9:1041034.
 33. Ding, K, Zhou M, Wang Z, Liu Q, Arnold CW, Zhang S, Metaxas DN. Graph convolutional networks for multi-modality medical imaging: Methods, architectures, and clinical applications. *arXiv preprint arXiv:2202.08916*. 2022.
 34. Zhang M, Long D, Chen Z, Fang C, Li Y, Huang P, Chen F, Sun H. Multi-view graph network learning framework for identification of major depressive disorder. *Comput Biol Med* 2023;166:107478.
 35. Al-Sabri R, Gao J, Chen J, Oloulade BM, Lyu T. Multi-View Graph Neural Architecture Search for Biomedical Entity and Relation Extraction. *IEEE/ACM Trans Comput Biol Bioinform* 2023;20:1221-33.
 36. Xiao S, Lin H, Wang C, Wang S, Rajapakse JC. Graph

- Neural Networks With Multiple Prior Knowledge for Multi-Omics Data Analysis. *IEEE J Biomed Health Inform* 2023;27:4591-600.
37. Fan K, Tang S, Gökbağ B, Cheng L, Li L. Multi-view graph convolutional network for cancer cell-specific synthetic lethality prediction. *Front Genet* 2022;13:1103092.
 38. Zhang Y, Zhou B, Song K, Sui X, Zhao G, Jiang N, Yuan X. PM2F2N: Patient Multi-view Multi-modal Feature Fusion Networks for Clinical Outcome Prediction. *Findings of the Association for Computational Linguistics: EMNLP 2022*:1985-94.
 39. Han Z, Chen F, Zhang H, Yang Z, Liu W, Shen Z, Xiong H. An attention-based representation learning model for multiple relational knowledge graph. *Expert Systems* 2023;40:e13234.
 40. Yang X, Zhang Y, Hu F, Deng Z, Zhang X. Feature aggregation-based multi-relational knowledge reasoning for COPD intelligent diagnosis. *Computers and Electrical Engineering* 2024;114:109068.
 41. Dai G, Wang X, Zou X, Liu C, Cen S. MRGAT: Multi-Relational Graph Attention Network for knowledge graph completion. *Neural Netw* 2022;154:234-45.
 42. Peng H, Zhang R, Dou Y, Yang R, Zhang J, Philip SY. Reinforced neighborhood selection guided multi-relational graph neural networks. *ACM Transactions on Information Systems (TOIS)* 2021;40:1-46.
 43. Baek J W, Chung K. Multi-Context Mining-Based Graph Neural Network for Predicting Emerging Health Risks. *IEEE Access* 2023;11:15153-63.
 44. D'Souza NS, Wang H, Giovannini A, Foncubierta-Rodriguez A, Beck KL, Boyko O, Syeda-Mahmood TF. Fusing modalities by multiplexed graph neural networks for outcome prediction from medical data and beyond. *Med Image Anal* 2024;93:103064.
 45. Jia X, Jiang M, Dong Y, Zhu F, Lin H, Xin Y, Chen H. Multimodal heterogeneous graph attention network. *Neural Computing and Applications* 2023;35:3357-72.
 46. Szegedy C, Ioffe S, Vanhoucke V, Alemi A. Inception-v4, inception-resnet and the impact of residual connections on learning. *Proceedings of the AAAI Conference on Artificial Intelligence* 2017.
 47. Han X, Yu Z, Zhuo Y, Zhao B, Ren Y, Lamm L, Xue X, Feng J, Marr C, Shan F, Peng T, Zhang XY. The value of longitudinal clinical data and paired CT scans in predicting the deterioration of COVID-19 revealed by an artificial intelligence system. *iScience* 2022;25:104227.
 48. Cui C, Yang H, Wang Y, Zhao S, Asad Z, Coburn LA, Wilson KT, Landman BA, Huo Y. Deep multimodal fusion of image and non-image data in disease diagnosis and prognosis: a review. *Prog Biomed Eng (Bristol)* 2023.
 49. Kim S, Lee N, Lee J, Hyun D, Park C. Heterogeneous Graph Learning for Multi-modal Medical Data Analysis. *Proceedings of the AAAI Conference on Artificial Intelligence* 2023;37:5141-50.
 50. Huang Z. Clustering large data sets with mixed numeric and categorical values. *Proceedings of the 1st pacific-asia conference on knowledge discovery and data mining, (PAKDD)* 1997:21-34.
 51. Pearson K. Contributions to the Mathematical Theory of Evolution. III. Regression, Heredity, and Panmixia. *Proceedings of the Royal Society of London* 1895;59:69-71.
 52. Bakr S, Gevaert O, Echegaray S, Ayers K, Zhou M, Shafiq M, Zheng H, Benson JA, Zhang W, Leung ANC, Kadoch M, Hoang CD, Shrager J, Quon A, Rubin DL, Plevritis SK, Napel S. A radiogenomic dataset of non-small cell lung cancer. *Sci Data* 2018;5:180202.
 53. Wang S, Shi J, Ye Z, Dong D, Yu D, Zhou M, Liu Y, Gevaert O, Wang K, Zhu Y, Zhou H, Liu Z, Tian J. Predicting EGFR mutation status in lung adenocarcinoma on computed tomography image using deep learning. *Eur Respir J* 2019;53:1800986.
 54. Xiong JF, Jia TY, Li XY, Yu W, Xu ZY, Cai XW, Fu L, Zhang J, Qin BJ, Fu XL, Zhao J. Identifying epidermal growth factor receptor mutation status in patients with lung adenocarcinoma by three-dimensional convolutional neural networks. *Br J Radiol* 2018;91:20180334.
 55. Jia L, Wu W, Hou G, Zhao J, Qiang Y, Zhang Y, Cai M. Residual neural network with mixed loss based on batch training technique for identification of EGFR mutation status in lung cancer. *Multimed Tools Appl* 2023. [Epub ahead of print]. doi: 10.1007/s11042-023-14876-2.
 56. Hu Z, Dong Y, Wang K, Sun Y. Heterogeneous graph transformer. *Proceedings of the Web Conference* 2020:2704-10.
 57. Lv Q, Ding M, Liu Q, Chen Y, Feng W, He S, Zhou C, Jiang J, Dong Y, Tang J. Are we really making much progress? Revisiting, benchmarking, and refining heterogeneous graph neural networks. *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining* 2021:1150-60.
 58. Parisot S, Ktena SI, Ferrante E, Lee M, Moreno RG, Glocker B, Rueckert D. Spectral graph convolutions for population-based disease prediction. *MICCAI 2017: 20th International Conference* 2017:177-85.
 59. Li S, Li WT, Wang W. Co-gcn for multi-view semi-

- supervised learning. Proceedings of the AAAI Conference on Artificial Intelligence 2020;34:4691-8.
60. Kong Y, Niu S, Gao H, Yue Y, Shu H, Xie C, Zhang Z, Yuan Y. Multi-stage graph fusion networks for major depressive disorder diagnosis. IEEE Transactions on Affective Computing 2022;13:1917-28.
 61. Mu W, Jiang L, Shi Y, Tunali I, Gray JE, Katsoulakis E, Tian J, Gillies RJ, Schabath MB. Non-invasive measurement of PD-L1 status and prediction of immunotherapy response using deep learning of PET/CT images. J Immunother Cancer 2021;9:e002118.
 62. Parimbelli E, Marini S, Sacchi L, Bellazzi R. Patient similarity for precision medicine: A systematic review. J Biomed Inform 2018;83:87-96.
 63. Brown SA. Patient Similarity: Emerging Concepts in Systems and Precision Medicine. Front Physiol 2016;7:561.
 64. Mao C, Yao L, Luo Y. ImageGCN: Multi-Relational Image Graph Convolutional Networks for Disease Identification With Chest X-Rays. IEEE Trans Med Imaging 2022;41:1990-2003.
 65. Dearden S, Stevens J, Wu YL, Blowers D. Mutation incidence and coincidence in non small-cell lung cancer: meta-analyses by ethnicity and histology (mutMap). Ann Oncol 2013;24:2371-6.
 66. Judd J, Abdel Karim N, Khan H, Naqash AR, Baca Y, Xiu J, et al. Characterization of KRAS Mutation Subtypes in Non-small Cell Lung Cancer. Mol Cancer Ther 2021;20:2577-84.
 67. NCCN. "NCCN clinical practice guidelines in oncology non-small cell lung Cancer (Version 3.2023–April 13,2022)". USA: NCCN 2023. Available online: <http://www.nccn.org/>
 68. Van der Maaten L, Hinton G. Visualizing data using t-SNE. J Mach Learn Res 2008;9:2579-605.

Cite this article as: Yang W, YOSHIDA S, Zhao J, Wu W, Qiang Y. MVASA-HGN: multi-view adaptive semantic-aware heterogeneous graph network for *KRAS* mutation status prediction. Quant Imaging Med Surg 2025;15(2):1190-1211. doi: 10.21037/qims-24-1370