Multi-Modal Learning for Predicting the Progression of Transarterial Chemoembolization Therapy in Hepatocellular Carcinoma

Lingzhi Tang[1,2,3], Haibo Shao[4], Jinzhu Yang[1,2,3(✉)], Jiachen Xu[1,2,3], Jiao Li[4], Yong Feng[1,2,3], Jiayuan Liu[1,2,3], Song Sun[1,2,3], and Qisen Wang[1,2,3]

[1] School of Computer Science and Engineering, Northeastern University, Shenyang, China
[2] Key Laboratory of Intelligent Computing in Medical Image, Ministry of Education, Northeastern University, Shenyang, China
[3] National Frontiers Science Center for Industrial Intelligence and Systems Optimization, Shenyang, China
yangjinzhu@cse.neu.edu.cn
[4] Department of Interventional Radiology, The First Hospital of China Medical University, Shenyang, China

Abstract. Hepatocellular carcinoma (HCC) is marked by high morbidity and is often diagnosed in middle or late stages. Transarterial chemoembolization (TACE) stands as the current standard of care for intermediate-stage HCC patients. Nevertheless, the tumor's heterogeneity significantly impacts patient prognosis. In this paper, a new dynamic multi-model graph network fusion multi-sequence magnetic resonance imaging is proposed to predict the prognosis of HCC patients after TACE treatment. The model proposes a spatial graph convolution module focusing on active regions within the tumor, a multi-module dynamic fusion module capturing the potential relationship between the tumor and the liver, and a cross-model topology fusion module using topological information to guide the multi-sequence MRI fusion. Our method achieved the best results compared to the state-of-the-art method, with an ACC of 75.27%, AUC of 76.69%, F1 of 73.84%, C-index of 0.6978, HR of 3.1988.

Keywords: Hepatocellular carcinoma · Progression-free survival · Multi-Modal Learning

1 Introduction

Hepatocellular carcinoma (HCC) ranks third in the world in terms of cancer-related mortality, and about 80% of HCC cases were diagnosed at an intermediate to advanced stage [2], depriving patients of surgical treatment and leaving

Lingzhi Tang and Haibo Shao contributed equally and can be considered co-first authors.

Z. Lin et al. (Eds.): PRCV 2024, LNCS 15044, pp. 178–193, 2025
https://doi.org/10.1007/978-981-97-8496-7_13

them with only options to prolong their lives, such as interventional and systemic therapies. Transarterial chemoembolization (TACE) is a common treatment for unresectable HCC and is effective in prolonging survival [4,13]. The heterogeneity of HCC may result in non-response to TACE treatment in certain patients [3,18], subsequently causing incomplete tumor necrosis and liver function deterioration. Therefore, it is crucial to develop accurate preoperative prediction models to identify patients who may benefit from TACE treatment.

Previous studies have mainly used preoperative CT [1] or preoperative magnetic resonance imaging (MRI) [42] radiological features combined with clinical features to construct prognostic models. Wang et al. [32] proposed a prediction model for post-TACE survival in patients with intermediate-stage hepatocellular carcinoma using CT images, while Liu [21] utilized radiomics to predict tumor response and prognosis post-TACE. The above-mentioned research mainly relies on the development of predefined manual features and cannot cope with the complex situations of each patient. Presently, deep learning algorithms possess the capacity to thoroughly analyze images from patients with HCC to extract rich and valuable information, thereby greatly benefiting clinical diagnosis and prognosis [8,38]. Convolutional neural networks (CNN), which are extensively employed in computer vision (CV), demonstrate high efficacy in predicting HCC prognosis [15,32]. However, the above studies simply use classical methods in CV, ignoring many problems:

Excessive introduction of irrelevant regions. CNN [14] and Transformers [11] treat images as grid or sequence structures, which are not flexible enough to deal with irregular and complex objects, especially tumors with irregular boundaries and complex spatial shapes. Existing visual graph neural networks (GNN) [12] deal with the above limitations by dividing the image into multiple patches and thus extracting the graph representation, but even this ignores the necrotic regions inside the tumor.

Lack of medical knowledge guidance. Previous studies have confirmed that non-tumor organ regions also influence cancer prognosis [36,39], but existing methods do not allow efficient fusion between organs and tumors, and secondly, HCC is usually multifocal, and potential relationships between multiple lesions cannot be extracted.

Inefficient fusion between multi-sequence features. Early multimodal fusion methods include channel concatenation [31] and feature extraction through multiple branches followed by splicing [34]. Subsequent approaches involve utilizing LSTM [40] or Transformer architectures [33] for feature extraction and integration, yet these methods rely solely on data-driven fusion processes without guidance from tumor topology information, leading to inefficient fusion, particularly in medical prognostic tasks with limited data.

Therefore, this study introduces a multi-modal graph convolutional fusion network, leveraging multi-sequence MRI, to overcome the aforementioned constraints and facilitate the accurate prediction of transarterial chemoembolization therapy outcomes for hepatocellular carcinoma. A new dynamic multi-model graph network (DMMGNet) was proposed containing three innovative blocks

such as spatial graph convolution module (SGCM), multi-module dynamic fusion module (MMDFM), and cross-model topology fusion module (CMTFM). Building upon the vision GNN proposed by [12], SGCM was proposed to dynamically fuse intra-tumour features to extract tumor spatial relationships. Furthermore, we propose MMDFM that can construct dynamic graphs using similar features of the tumor and liver to achieve efficient feature interactions thereby capturing potential prognostic representations. Finally, CMTFM is proposed to use topological relationship guidance for fusion of intra-sequence features in multi-sequence MRI The contributions and innovations of this paper are summarised below:

- We developed a DMMGNet that combines information from multi-sequence MRI to predict the prognosis of HCC patients undergoing TACE.
- In this study, we proposed SGCM for extracting spatial information of tumors, MMDFM for capturing the potential relationship between the liver and tumors, and CMTFM for utilizing tumor topology to guide the fusion of multi-sequence MRI features.
- Our model achieves an accuracy (ACC) of 0.75, an area under the curve (AUC) of 0.77, a concordance index (C-Index) of 0.70, and a hazard ratio (HR) of 3.2, representing a significant improvement compared to existing methods.

The paper is structured as follows: Sect. 2 provides a detailed introduction to our model architecture, Sect. 3 presents a summary of the data preparation stage along with the experimental results and finally, Sect. 4 contains the discussions and conclusion.

2 Related Works

Early HCC prognostic methods [10,16,27] were based on staging systems such as the Barcelona Clinic Liver Cancer staging system, the Hong Kong Liver Cancer staging system, and the Pre-TACE Predict model. These models relied on limited features, resulting in poor performance on external test sets [17,24]. With advances in image processing and machine learning, researchers now use image-based methods like radiomics and AI, supported by widespread medical imaging, to predict HCC treatment response by extracting higher-level latent features. CT is a commonly used medical imaging for extracting numerous radiomic features and employing machine learning methods to build models predicting the initial response to TACE treatment [25,26]. In recent years, some studies have begun developing prognostic models based on ultrasound radiomics [20,23,35], achieving certain progress. Multi-sequence MRI offers the advantage of providing rich soft tissue contrast and diverse physiological information. However, it is limited by higher acquisition difficulty and under-explored multimodal prognostic fusion methods, resulting in insufficient research progress.

In recent years, multimodal prediction has increasingly because of its important potential in tumor prognostic assessment. The earliest and most common

methods include simply concatenating multiple modality features of MRI inputs [31], or employing multi-branch approaches to extract features from different modalities separately [34]. Building upon these methods, Zhan et al. [41] and Wang et al. [33] proposed early cancer recurrence prediction methods based on transformer fusion. However, the introduction of numerous parameters by transformers poses challenges for training on small datasets. To address these challenges, Gao et al. [9] designed a CNN-RNN classifier that effectively integrates diagnostic imaging and clinical features of liver cancer patients, demonstrating improved training outcomes. Yao et al. [37] utilized multi-branch methods to extract prognostic features at different levels, successfully applied to pancreatic cancer prognosis. Tang et al. [29] introduced a novel approach for predicting colorectal cancer liver metastasis prognosis, exploring new pathways in multi-sequence MRI feature fusion. However, the fusion process lacks efficient feature interaction and confident guidance on topological methods.

3 Method

In this section, we will first briefly introduce the main structure of the paper, and subsequently we will detail the proposed SGCM, MMDFM, CMTFM, and finally the prognostic task.

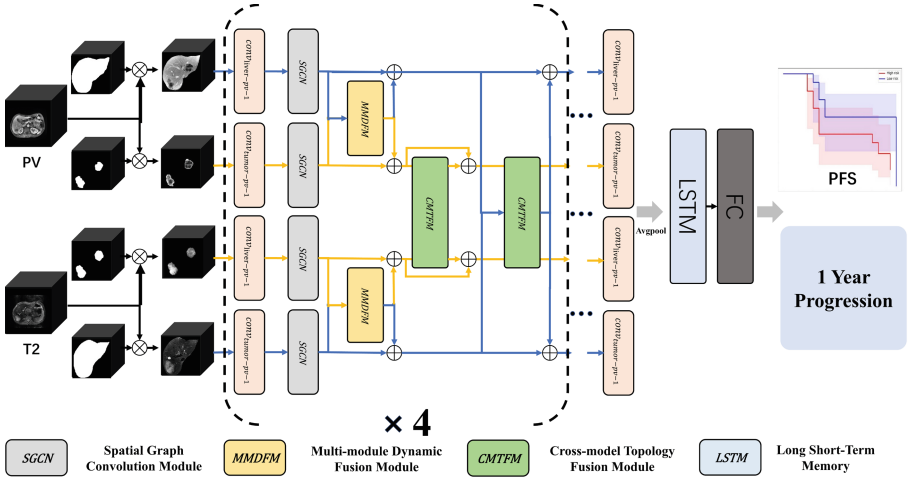

Fig. 1. Our proposed DMMGNet.

3.1 Main Architecture of DMMGNet

The illustration of our proposed network is shown in Fig. 1, where the main body has four parallel convolutional branches dealing with liver and tumor for T2-weighted and PV[1] sequences, respectively (liver branch in blue arrow lines and tumor branch in yellow arrow lines in Fig .1). The convolutional layer in each branch is followed by SGCN designed to extract the tumor spatial relationship, MMDFM is designed to extract the potential representation for the liver and tumor branches between each sequence, and CMTFM is designed for inter-sequence feature-to-feature interaction for the corresponding branches between sequences. The convolutional operations of the liver and tumor branches in the T2 sequence are denoted as $conv_{liver-i-T2}, conv_{tumor-i-T2}, i = 1, 2, 3, 4$, respectively. The output feature maps at each stage are represented as $FLT2_i \in R^{d_i \times w_i \times h_i \times c_i}$, $FTT2_i \in R^{d_i \times w_i \times h_i \times c_i}$, and the corresponding mask maps are also denoted as $MLT2_i \in R^{d_i \times w_i \times h_i \times 1}$, and $MTT2_i \in R^{d_i \times w_i \times h_i \times 1}$, $i = 1, 2, 3, 4$ which can be summarized in below:

$$FTT2_i = conv_{tumor-i-t2}\left(FTT2_{i-1}\right) \tag{1}$$

$$FTT2_i = SGCM\left(FTT2_i\right) \tag{2}$$

$$FTT2_i, FLT2_i = MMDFM(FTT2_i, FLT2_i) \tag{3}$$

$$FTPV_i, FTT2_i = CMTFM(FTPV_i, FTT2_i, MTPV_i, MLPV_i) \tag{4}$$

The PV sequence branch also has the same operation and the corresponding feature map output as $FLPV_i \in R^{d_i \times w_i \times h_i \times c_i}$, and $FTPV_i \in R^{d_i \times w_i \times h_i \times c_i}$.

3.2 Spatial Graph Convolution Module

The irregular shapes of the tumor and liver regions pose a challenge to the analysis of standard convolutional operations as they introduce many uncorrelated regions. Inspired by the local feature fusion mechanism [28], we introduce dynamic node selection based on GCN [19] to simulate the tumor interior active regions undergoing feature interactions.

As shown in Fig. 2, the SGCM consists of three main parts, node selection, feature interaction, and feature local replacement.

Node selection Each $f_i \in w \times h \times c$ tensor in an image feature map F is obtained by summing over the channel dimensions to get $index_i \in w \times h$, representing the high-level compressed representation corresponding to the w*h positions. Based on $index$, the top-k algorithm is used to select the largest k positions and gather the local feature representations $\phi \in k \times c$ of the feature map F corresponding to these positions.

$$\phi_i = gather\left(top_k(index_i), f_i\right), i \in 1 \cdots d \tag{5}$$

[1] For simplicity, we abbreviate the contrast-enhanced T1-weighted imaging (CE-T1WI) portal vein phase as PV.

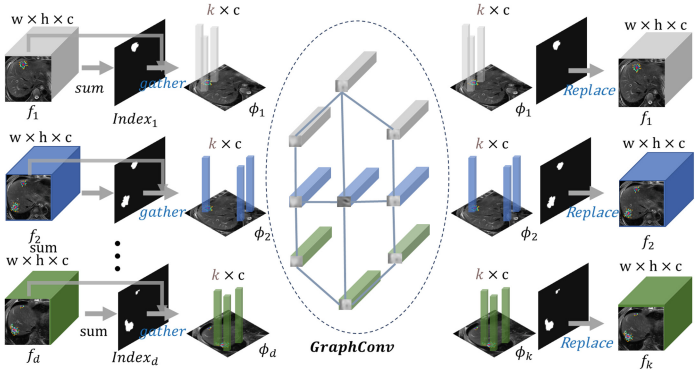

Fig. 2. Spatial Graph Convolution Module.

All local features taken from the d-layer feature map are used for subsequent graph network construction and feature interaction.

Feature interaction To adapt to the diversity of patients' tumors, we choose to adopt the GCN framework of [19], whose main advantages are as follows:

- The Dilated k-NN function was used to construct dynamic graph structures, thus adapting to the micro-environment inside different patient tumors.
- The aggregation process uses a random expansion to select neighbors for feature interaction and improves generalization.

Feature local replacement After feature interaction, we replace the interacted local features to the original position according to the previously recorded *index*.

3.3 Multi-module Dynamic Fusion Module

The Multi-module Dynamic Fusion Module is designed to fuse liver and tumor features to capture potential prognostic associations. The liver region is much larger than the tumor region, and direct fusion will inevitably introduce too many irrelevant regions leading to inefficient fusion. Inspired by [29], we introduced a dynamic feature filtering process based on the SGCM proposed in Sect. 3.2, as shown in Fig. 3.

After node selection, we take the obtained local feature representations $(tumor_1, \ldots, tumor_d)$ concatenated into $Tumor \in k * d \times C$, and similarly $Liver \in k * d \times C$. Subsequently, the following formula is used to calculate the similarity between each pair of positions in the tumor and liver to construct the similarity map M

$$M^{(i,j)} = g(Tumor^{(i)}, Liver^{(j)}) \tag{6}$$

where $i \in \{1, \ldots, k \times d\}, j \in \{1, \ldots, k \times d\}$ and g is a similarity metric. Using the top-k algorithm on different dimensions of the graph M, *liver_sim_index*

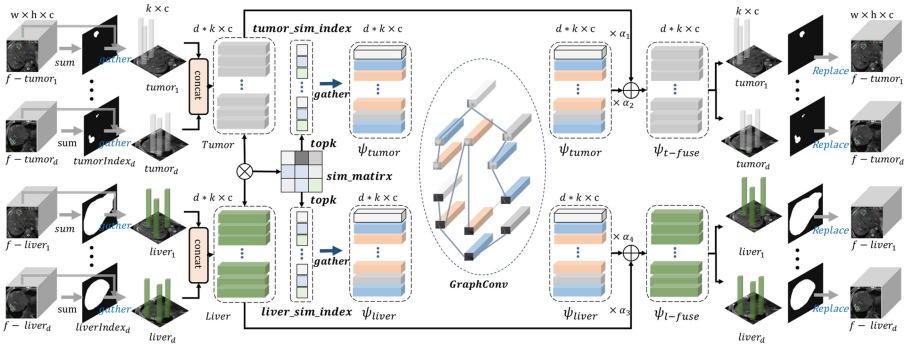

Fig. 3. Multi-module Dynamic Fusion Module

and *liver_sim_index* are obtained respectively. These indices were then used to collect important liver features and tumor features to obtain $\Psi_{liver} \in R^{d*k\times c}$ and $\Psi_{tumor} \in R^{d*k\times c}$. The subsequent operations are consistent with SGCM, using GCN for feature interaction, the only difference is that we introduce an adaptive coefficient vector $alpha = [\alpha_1, \ldots, \alpha_4]$ to select each position. The fusion process is represented as:

$$\Phi_{t-fuse} = \alpha_1 \cdot Tumor + \alpha_2 \cdot \Psi_{tumor} \tag{7}$$

$$\Phi_{l-fuse} = \alpha_1 \cdot Liver + \alpha_2 \cdot \Psi_{liver} \tag{8}$$

Finally, the fuses the local representations Φ_{fuse} to their corresponding positions to obtain the new feature mapping.

3.4 Cross-model Topology Fusion Module

In the diagnostic process of doctors, they often observe tissues at the same location of different sequences to determine whether it is a tumor or the boundary of a tumor, which not only reduces the misclassification caused by a single sequence (e.g., cysts) but also suggests that the spatial topological relationship between multiple sequences may help feature fusion. Figure 4 illustrates the details of the proposed CMTFM. For instance, the inputs for the CMTFM include $FTPV$, $FTT2$, and their respective masks $MTPV$ and $MTT2$. Firstly, $MTPV$ is decomposed into $3*3*3$ patches, finding the largest patch as the center patch of the tumor (the small red cube in Fig. 4). The adjacent patches in 26 directions at the center are regarded as first-order neighbors, and the neighbors one level further out from the first-order neighbors are regarded as second-order neighbors, finally, we get ϕ_{PV} and ϕ_{T2}. Feature fusion is performed using GCN for the first-order neighbors of the two sequences and the same is also done for the features of the second-order neighbors. Finally, the fused features are substituted back to the initial position of the original feature map.

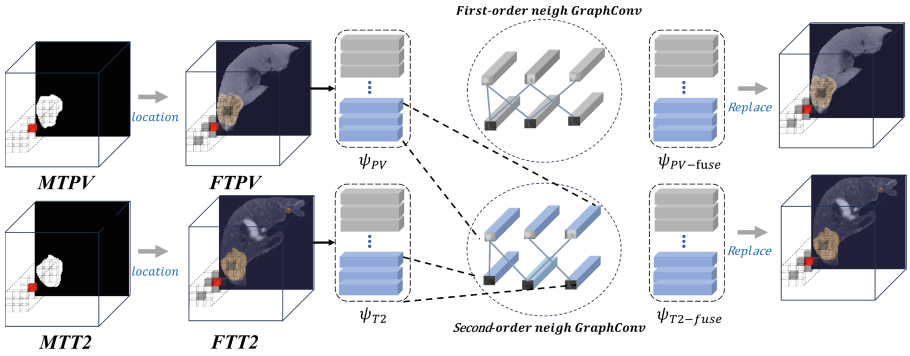

Fig. 4. Cross-model Topology Fusion Module

4 Experiments and Results

4.1 Dataset Description and Implementation Details

The dataset collected pre-treatment MRI and clinically relevant data from 102 patients who underwent transarterial chemoembolization and targeted immunotherapy from January 2019 to December 2022 at the First Hospital of China Medical University, marked by two experienced radiologists, and was approved under ethical number AF-SOP-07-1.2-01. Each patient's input to the network comprises 40 images sized 112x112 pixels, including 10 consecutive slices of T2-weighted tumor images, 10 slices of T2-weighted liver images, 10 slices of PV tumor images, and 10 slices of PV liver images.

4-fold cross-validation is performed to evaluate the performance of our proposed method. All experiments were performed using PyTorch 1.6.0, utilizing 4 NVIDIA A100 GPUs and an Intel(R) Xeon(R) Silver 4314 CPU @ 2.40GHz. For optimization, we utilized an Adam optimizer with a learning rate of 5e-3 and a batch size of 16. The learning rate was initialized at 0.005 and decayed by a factor of 0.8 in case the loss did not decrease during training.

Two tasks were devised to validate the network's performance: one involves predicting 1-year Progression, while the other pertains to Progression-free survival(PFS). The loss function employed is the negative log partial likelihood [7], which incorporates the predicted risk (r) and censored data (E), and can be computed utilizing the partial Cox model:

$$Loss_{neg} = -\frac{1}{\sum_i E_i} \sum_{i=1}^{n} E_i \left(r_i - log \sum_{j \in R_i} \exp(r_j) \right) \tag{9}$$

Here, R_i represents the cohort of individuals who experience the event at or after the observed time for the i^{th} sample. For the classification experiments, we employ three evaluation metrics: Accuracy (ACC), Area Under Curve (AUC), and F1 Score (F1). In the ranking experiment, we utilize the Concordance index (C-index) and Hazard ratio (HR).

4.2 Ablation Study

This section evaluates the efficacy of each module in the proposed DMMGNet. To validate the role of each module individually, four architectural variants (Baseline, Baseline+SGCM, Baseline+MMDFM, Baseline+CMTFM) were trained using consistent hyperparameters. The baseline architecture comprises four parallel CNN branches, each processing a distinct input. Subsequently, the outputs of these branches are equally pooled and fed into an LSTM. Other variants of the model include additional modules appended to the Baseline architecture. As shown in Table 1, each module has improved in various metrics compared to Baseline, with CMTFM improving by 3.68% in ACC, SGCM by 2.33% in AUC, MMDFM by 1.9 % in C-index, and CMTFM by 0.60 in HR.

Table 1. Ablation studies for different design choices. Bold indicates the best performance.

No	Components				1-year Progression			PFS	
	Baseline	SGCM	MMDFM	CMTFM	ACC	AUC	F1	C-index	HR
1	✓				0.6632±0.0632	0.7116±0.0072	0.6334±0.0424	0.6499±0.0327	2.2280±0.8027
2	✓	✓			0.6685±0.0652	0.7349±0.0471	0.6622±0.0178	0.6659±0.0739	2.3405±0.4849
3	✓		✓		0.6737±0.0173	0.7183±0.0441	0.6622±0.0179	0.6689±0.0422	2.3361±0.0513
4	✓			✓	0.7000±0.0346	0.7216±0.0119	0.6855±0.0505	0.6573±0.0520	2.8243±0.2628
5	✓	✓	✓		0.7085±0.0460	0.7283±0.0322	0.7360±0.0371	0.6706±0.0204	2.6953±1.1178
6	✓	✓		✓	0.7025±0.0334	0.7316±0.0240	0.6957±0.0356	0.6797±0.0153	2.8704±1.2616
7	✓		✓	✓	0.7127±0.0194	0.7293±0.0110	0.7038±0.0154	0.6723±0.0342	3.0409±1.1644
8	✓	✓	✓	✓	**0.7527±0.0127**	**0.7669±0.0335**	**0.7384±0.0258**	**0.6978±0.0127**	**3.1988±0.6413**

Then, we investigated the complementary effect between modules by combining the proposed modules two by two to compare with the previous experiments (SGCM+MMDFM, SGCM+CMTFM, MMDFM+CMTFM). The results demonstrate that the modules complement each other, leading to improvements in all metrics. Specifically, the F1 score of SGCM+MMDFM reached 73.6%, the AUC of SGCM+CMTFM achieved 73.16%, and the Hazard Ratio of MMDFM+CMTFM reached 3.0479.

Finally, the three modules were integrated to create the final DMMGNnet, yielding optimal results: ACC of 75.27%, AUC of 76.69%, F1 of 73.84%, C-index of 0.6978, and HR of 3.1988. As shown in Fig. 5, the AUC is improved by 5.53% compared to the initial baseline AUC. The 4-fold Kaplan-Meier analysis shown in Fig. 6 shows that only DMMGNet performs well in the 4-fold assessment, whereas the other methods with unstable high-risk and low-risk curves show an overlapping lack of generalization.

4.3 Comparisons with the SOTA methods

We replicated several recently developed deep learning prediction methods [5,6,22,29,30,34,37]. We standardized input image size, epochs, batch size, and the division of training and validation sets to ensure a fair comparison of deep

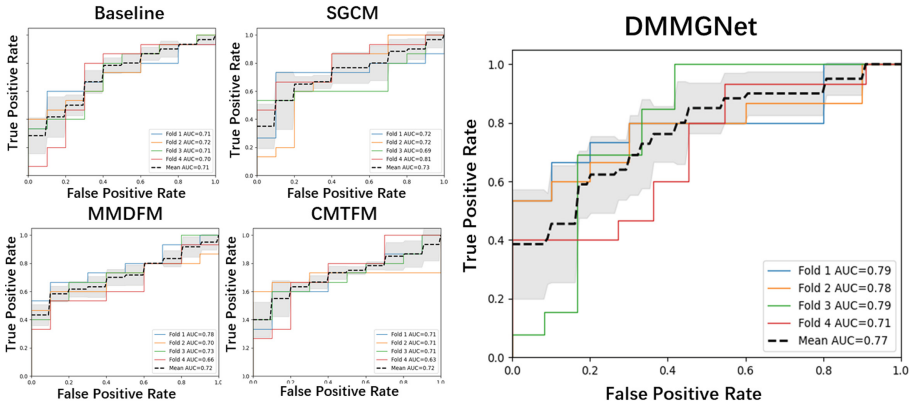

Fig. 5. The Receiver Operating Characteristic (ROC) curves for five architectures are depicted. The black dashed line indicates the average ROC, the solid line corresponds to the result of the 4-fold test, and the shaded area signifies the variance's addition and subtraction.

learning techniques, consistent with the settings in Sect. 4.1. All image-based methods in the training dataset underwent equal preprocessing and augmentation.

Table 2. Comparison with SOTA in the HCC dataset. Bold indicates the best performance.

HCC	1-year Progression			PFS	
	ACC	AUC	F1	C-index	HR
Chen et al.[6]	0.6856±0.0535	0.5797±0.0673	0.6343±0.0355	0.5632±0.0342	1.7900±0.6864
Chen et al.[5]	0.7031±0.0535	0.6089±0.0478	0.6585±0.0355	0.6180±0.0317	2.0527±0.4698
Liu et al.[22]	0.6567±0.0189	0.6589±0.0499	0.6550±0.0601	0.6102±0.0634	2.0071±0.5372
Tang et al.[30]	0.5921±0.0502	0.5765±0.0365	0.5358±0.0357	0.5851±0.0451	1.8920±0.6754
Yao et al.[37]	0.6038±0.1098	0.5860±0.1159	0.6657±0.1267	0.5755±0.0581	2.1402±0.7611
Wang et al.[34]	0.6367±0.0367	0.6389±0.0520	0.6017±0.0601	0.6022±0.0457	2.0100±0.7802
Tang et al.[29]	0.7427±0.0334	0.7450±0.0340	0.7234±0.0284	0.6838±0.0398	2.7950±0.4293
Ours	**0.7527±0.0127**	**0.7669±0.0335**	**0.7384±0.0258**	**0.6987±0.0127**	**3.1988±0.6413**

The results in Table 2 demonstrate the superiority of our method in performing these two tasks. Specifically, in the task of predicting PFS, the HR exceeds the next best by 0.4038, and the C-index exceeds the next best by 0.0149. In the classification task of 1-year progression, the ACC exceeds 1%, the AUC exceeds 2.19%, and the F1 score exceeds 1.5%.

5 Discussions and Conclusions

Hepatocellular carcinoma has a high incidence rate and is often diagnosed late. Transcatheter arterial chemoembolization is currently the standard treatment

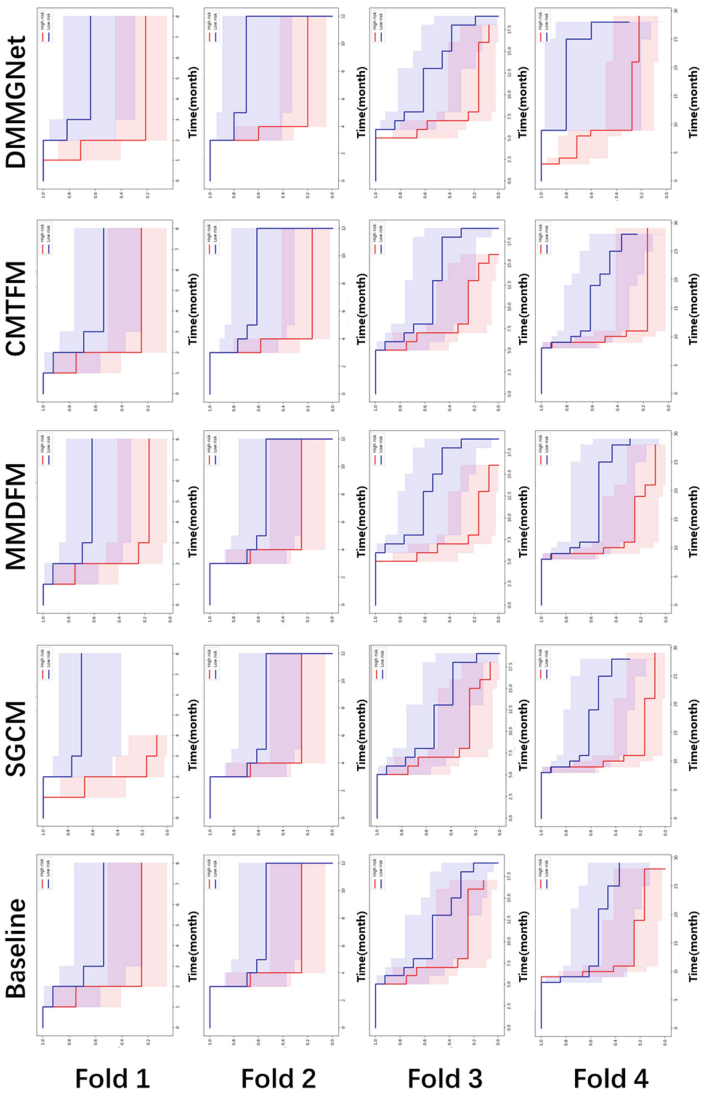

Fig. 6. The four-fold Kaplan-Meier analyses of the five structures, each row represents the result of one structure, and each column represents the result of one fold. The blue line represents low risk, while the red line represents high risk. The greater the separation on between the two lines, the greater the difference in risk between the two groups.

for intermediate-stage HCC patients and is widely used. Tumor heterogeneity affects patient prognosis, and deep learning provides a new avenue for precise and personalized treatment of tumors. This study utilizes deep learning based on pre-operative multi-sequence MRI to predict 1-year progression and progression-free survival in patients with intermediate-stage HCC undergoing TACE combined with targeted therapy and immune therapy. We proposed a new DMMGNet that achieved the best results ACC of 75.27%, AUC of 76.69%, F1 of 73.84%, c-index of 0.6978,HR of 3.1988.

Previous studies have shown a lack of specificity and generalizability when dealing with HCC prognosis prediction tasks. These studies either relied on predefined traditional histological features [21,32] or resorted to generic deep learning networks [15]. However, these approaches often fail to achieve sufficient and accurate generalization for data-constrained specific tasks, especially during multi-fold cross-validation, with poorer results and larger variance observed in Table 2. Recent methods proposed for prognostic prediction of rectal cancer liver metastasis resection [29] started to focus on how medical prior knowledge can be fused into the network, complementary multi-sequence MRI, etc. Thus their method is closest in effect to ours, achieving suboptimal results.

Cross-disciplinary tasks are the mainstream of the times, unlike the big data-driven natural image field, we believe that a combination of medical imaging and deep learning technology is the expression of medical a priori knowledge in neural networks, which is also the main contribution of this paper. The innovation of DMMGNet lies in its proposal of three innovative modules. For ease of extension, the input and output sizes of these three modules are consistent and summarized into three components: feature selection, feature fusion, and feature replacement, with the main difference lying in feature selection. Tumors typically comprise necrotic regions and tumor cells, with the latter often having a greater impact on prognosis. In light of this, SGCM aggregates over channels to obtain indices of significant spatial locations, enabling feature selection guided by medical prior knowledge. MMDFM aims to select meaningful features from the liver and tumor for feature interaction. We employ similarity measurement to facilitate this selection, thus avoiding the introduction of excessive irrelevant regions. By utilizing topological information to guide the fusion of multi-sequence features, CMTFM aligns with the prior knowledge of doctors, who often observe the same location across different sequences.

Despite achieving favorable outcomes in predicting HCC prognosis, there is still room for improvement and further research. Our envisioned scenario involves employing an automated segmentation model for tumor delineation to mitigate inter-observer and intra-observer variability and to automate the analysis pipeline. Given the scarcity of patient samples, this work faces challenges in data collection. Future investigations will focus on refining end-to-end efficient classification methods and enhancing the model's generalizability across multiple centers and large datasets. Such improvements in the model will lead to more accurate prognostic predictions for HCC, with greater versatility and ease of implementation in hospitals. In addition, the validation of the usage scenarios

in this study is limited. To our knowledge, there is currently no publicly available multi-sequence MRI prognostic data. Such data often involve medical ethics and require long-term follow-up for collection. In the future, we will validate the application of this method in the prognosis of multiple diseases to demonstrate its generalizability.

In summary, this paper introduces a novel HCC prognosis prediction network called DMMGNet. Rooted in medical prior knowledge representation, this network proposes SGCM to focus on the active regions within tumors, MMDFM to capture the potential relationship between tumors and the liver, and MMDFM utilizes topological information to guide fusion. Compared to the state-of-the-art methods, it achieves the best results, contributing to precise patient treatment and exploring potential prognosis mechanisms for HCC.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments. This work is supported in part by the National Natural Science Foundation of China, No. U22A2022; 111 Project (B16009)

References

1. Bai, H., Meng, S., Xiong, C., Liu, Z., Shi, W., Ren, Q., Xia, W., Zhao, X., Jian, J., Song, Y., et al.: Preoperative cect-based radiomic signature for predicting the response of transarterial chemoembolization (tace) therapy in hepatocellular carcinoma. Cardiovasc. Intervent. Radiol. **45**(10), 1524–1533 (2022)
2. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., Jemal, A.: Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. **68**(6), 394–424 (2018). https://doi.org/10.3322/caac.21492
3. Bruix, J., Reig, M., Sherman, M.: Evidence-based diagnosis, staging, and treatment of patients with hepatocellular carcinoma. Gastroenterology **150**(4), 835–853 (2016)
4. Bruix, J., Sala, M., Llovet, J.M.: Chemoembolization for hepatocellular carcinoma. Gastroenterology **127**(5), S179–S188 (2004)
5. Chen, J., Cheung, H.M., Milot, L., Martel, A.L.: Aminn: Autoencoder-based multiple instance neural network improves outcome prediction in multifocal liver metastases, pp. 752–761 (2021)
6. Chen, J., et al.: Unsupervised clustering of quantitative imaging phenotypes using autoencoder and gaussian mixture model, pp. 575–582 (2019)
7. Cheng, N.M., Yao, J., Cai, J., Ye, X., Zhao, S., Zhao, K., Zhou, W., Nogues, I., Huo, Y., Liao, C.T., et al.: Deep learning for fully automated prediction of overall survival in patients with oropharyngeal cancer using fdg-pet imaging. Clin. Cancer Res. **27**(14), 3948–3959 (2021)
8. Fu, S., Lai, H., Li, Q., Liu, Y., Zhang, J., Huang, J., Chen, X., Duan, C., Li, X., Wang, T., et al.: Multi-task deep learning network to predict future macrovascular invasion in hepatocellular carcinoma. EClinicalMedicine **42** (2021)

9. Gao, R., et al.: Deep learning for differential diagnosis of malignant hepatic tumors based on multi-phase contrast-enhanced ct and clinical data. J. Hematology Oncology **14**(1), 1–7 (2021)
10. Han, G., Berhane, S., Toyoda, H., Bettinger, D., Elshaarawy, O., Chan, A.W., Kirstein, M., Mosconi, C., Hucke, F., Palmer, D., et al.: Prediction of survival among patients receiving transarterial chemoembolization for hepatocellular carcinoma: a response-based approach. Hepatology **72**(1), 198–212 (2020)
11. Han, K., Wang, Y., Chen, H., Chen, X., Guo, J., Liu, Z., Tang, Y., Xiao, A., Xu, C., Xu, Y., et al.: A survey on vision transformer. IEEE Trans. Pattern Anal. Mach. Intell. **45**(1), 87–110 (2022)
12. Han, K., Wang, Y., Guo, J., Tang, Y., Wu, E.: Vision gnn: an image is worth graph of nodes. Adv. Neural. Inf. Process. Syst. **35**, 8291–8303 (2022)
13. Haywood, N., Gennaro, K., Obert, J., Sauer, P.F., Redden, D.T., Zarzour, J., Smith, J.K., Bolus, D., Saddekni, S., Aal, A.K.A., et al.: Does the degree of hepatocellular carcinoma tumor necrosis following transarterial chemoembolization impact patient survival? J. Oncology **2016** (2016)
14. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
15. Hsieh, C., Laguna, A., Ikeda, I., Maxwell, A.W., Chapiro, J., Nadolski, G., Jiao, Z., Bai, H.X.: Using machine learning to predict response to image-guided therapies for hepatocellular carcinoma. Radiology **309**(2), e222891 (2023)
16. O Kadalayil, L., Benini, R., Pallan, L., O'beirne, J., Marelli, L., Yu, D., Hackshaw, A., Fox, R., Johnson, P., Burroughs, A., et al.: A simple prognostic scoring system for patients receiving transarterial embolisation for hepatocellular cancer. Annals Oncology **24**(10), 2565–2570 (2013)
17. Kim, D.S., Kim, B.K., Lee, J.S., Lee, H.W., Park, J.Y., Kim, D.Y., Ahn, S.H., Kim, S.U.: Validation of pre-/post-tace-predict models among patients with hepatocellular carcinoma receiving transarterial chemoembolization. Cancers **14**(1), 67 (2021)
18. Kudo, M.: Proposal of primary endpoints for tace combination trials with systemic therapy: lessons learned from 5 negative trials and the positive tactics trial. Liver cancer **7**(3), 225–234 (2018)
19. Li, G., Muller, M., Thabet, A., Ghanem, B.: Deepgcns: Can gcns go as deep as cnns? In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (2019)
20. Liu, F., Liu, D., Wang, K., Xie, X., Su, L., Kuang, M., Huang, G., Peng, B., Wang, Y., Lin, M., et al.: Deep learning radiomics based on contrast-enhanced ultrasound might optimize curative treatments for very-early or early-stage hepatocellular carcinoma patients. Liver Cancer **9**(4), 397–413 (2020)
21. Liu, Q.P., Yang, K.L., Xu, X., Liu, X.S., Qu, J.R., Zhang, Y.D.: Radiomics analysis of pretreatment mri in predicting tumor response and outcome in hepatocellular carcinoma with transarterial chemoembolization: a two-center collaborative study. Abdominal Radiology, pp. 1–13 (2022)
22. Liu, Z., Sun, Q., Bai, H., Liang, C., Chen, Y., Li, Z.C.: 3d deep attention network for survival prediction from magnetic resonance images in glioblastoma. In: 2019 IEEE International Conference on Image Processing (ICIP), pp. 1381–1384 (2019). https://doi.org/10.1109/ICIP.2019.8803077
23. Ma, Q.P., He, X.l., Li, K., Wang, J.f., Zeng, Q.J., Xu, E.J., He, X.q., Li, S.y., Kun, W., Zheng, R.Q., et al.: Dynamic contrast-enhanced ultrasound radiomics for

hepatocellular carcinoma recurrence prediction after thermal ablation. Molecular Imaging Biol. **23**, 572–585 (2021)
24. Mähringer-Kunz, A., Kloeckner, R., Pitton, M.B., Düber, C., Schmidtmann, I., Galle, P.R., Koch, S., Weinmann, A.: Validation of the risk prediction models state-score and start-strategy to guide tace treatment in patients with hepatocellular carcinoma. Cardiovasc. Intervent. Radiol. **40**, 1017–1025 (2017)
25. Morshid, A., Elsayes, K.M., Khalaf, A.M., Elmohr, M.M., Yu, J., Kaseb, A.O., Hassan, M., Mahvash, A., Wang, Z., Hazle, J.D., et al.: A machine learning model to predict hepatocellular carcinoma response to transcatheter arterial chemoembolization. Radiol. Artif. Intell. **1**(5), e180021 (2019)
26. Peng, J., Huang, J., Huang, G., Zhang, J.: Predicting the initial treatment response to transarterial chemoembolization in intermediate-stage hepatocellular carcinoma by the integration of radiomics and deep learning. Front. Oncol. **11**, 730282 (2021)
27. Pinato, D.J., Arizumi, T., Allara, E., Jang, J.W., Smirne, C., Kim, Y.W., Kudo, M., Pirisi, M., Sharma, R.: Validation of the hepatoma arterial embolization prognostic score in european and asian populations and proposed modification. Clin. Gastroenterol. Hepatol. **13**(6), 1204–1208 (2015)
28. Tang, L., Wang, X., Yang, J., Wang, Y., Qu, M., Li, H.: Dlffnet: a new dynamical local feature fusion network for automatic aortic valve calcification recognition using echocardiography. Comput. Methods Programs Biomed. **243**, 107882 (2024)
29. Tang, L., Zhang, Z., Yang, J., Feng, Y., Sun, S., Liu, B., Ma, J., Liu, J., Shao, H.: A new automated prognostic prediction method based on multi-sequence magnetic resonance imaging for hepatic resection of colorectal cancer liver metastases. IEEE J. Biomed. Health Inform. **28**(3), 1528–1539 (2024). https://doi.org/10.1109/JBHI.2024.3350247
30. Tang, Z., Xu, Y., Jiao, Z., Lu, J., Jin, L., Aibaidula, A., Wu, J., Wang, Q., Zhang, H., Shen, D.: Pre-operative overall survival time prediction for glioblastoma patients using deep learning on both imaging phenotype and genotype. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part I 22, pp. 415–422. Springer (2019)
31. Tang, Z., Xu, Y., Jin, L., Aibaidula, A., Lu, J., Jiao, Z., Wu, J., Zhang, H., Shen, D.: Deep learning of imaging phenotype and genotype for predicting overall survival time of glioblastoma patients. IEEE Trans. Med. Imaging **39**(6), 2100–2109 (2020)
32. Wang, H., Liu, Y., Xu, N., Sun, Y., Fu, S., Wu, Y., Liu, C., Cui, L., Liu, Z., Chang, Z., et al.: Development and validation of a deep learning model for survival prognosis of transcatheter arterial chemoembolization in patients with intermediate-stage hepatocellular carcinoma. Eur. J. Radiol. **156**, 110527 (2022)
33. Wang, J., Mao, Y., Gao, X., Zhang, Y.: Recurrence risk stratification for locally advanced cervical cancer using multi-modality transformer network. Front. Oncol. **13**, 1100087 (2023)
34. Wang, W., Wang, F., Yang, Y., Li, Y., Liu, J., Han, X., Lin, L., Tong, R., Hu, H., Chen, Y.W.: Deep learning-based risk prediction model for recurrence-free survival in patients with hepatocellular carcinoma using multi-phase ct image. In: 2022 IEEE 11th Global Conference on Consumer Electronics (GCCE), pp. 926–929. IEEE (2022)
35. Wu, J.p., Ding, W.z., Wang, Y.l., Liu, S., Zhang, X.q., Yang, Q., Cai, W.j., Yu, X.l., Liu, F.y., Kong, D., et al.: Radiomics analysis of ultrasound to predict recurrence of hepatocellular carcinoma after microwave ablation. Int. J. Hyperthermia **39**(1), 595–604 (2022)

36. Wu, T.H., Hatano, E., Yamanaka, K., Seo, S., Taura, K., Yasuchika, K., Fujimoto, Y., Nitta, T., Mizumoto, M., Mori, A., et al.: A non-smooth tumor margin on preoperative imaging predicts microvascular invasion of hepatocellular carcinoma. Surg. Today **46**, 1275–1281 (2016)
37. Yao, J., et al.: Deepprognosis: preoperative prediction of pancreatic cancer survival and surgical margin via comprehensive understanding of dynamic contrast-enhanced ct imaging and tumor-vascular contact parsing. Med. Image Anal. **73**, 102150 (2021)
38. Yasaka, K., Akai, H., Abe, O., Kiryu, S.: Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced ct: a preliminary study. Radiology **286**(3), 887–896 (2018)
39. Yoneda, N., Matsui, O., Kobayashi, S., Kitao, A., Kozaka, K., Inoue, D., Yoshida, K., Minami, T., Koda, W., Gabata, T.: Current status of imaging biomarkers predicting the biological nature of hepatocellular carcinoma. Jpn. J. Radiol. **37**, 191–208 (2019)
40. Zhan, G., Wang, F., Wang, W., Li, Y., Chen, Q., Hu, H., Chen, Y.W.: A transformer-based model for preoperative early recurrence prediction of hepatocellular carcinoma with muti-modality mri. In: Asian Conference on Computer Vision, pp. 185–194. Springer (2022)
41. Zhan, G., et al.: A transformer-based model for preoperative early recurrence prediction of hepatocellular carcinoma with muti-modality mri. In: Computer Vision–ACCV 2022 Workshops: 16th Asian Conference on Computer Vision, Macao, China, December 4–8, 2022, Revised Selected Papers, pp. 185–194. Springer (2023)
42. Zhao, Y., Huang, F., Liu, S., Jian, L., Xia, X., Lin, H., Liu, J.: Prediction of therapeutic response of unresectable hepatocellular carcinoma to hepatic arterial infusion chemotherapy based on pretherapeutic mri radiomics and albumin-bilirubin score. J. Cancer Res. Clin. Oncol. **149**(8), 5181–5192 (2023)