MGNN: A Multimodal Graph Neural Network for Predicting the Survival of Cancer Patients

Jianliang Gao
School of Computer Science and Engineering, Central South University

Tengfei Lyu
School of Computer Science and Engineering, Central South University

Fan Xiong
School of Computer Science and Engineering, Central South University

Jianxin Wang
School of Computer Science and Engineering, Central South University

Weimao Ke
College of Computing & Informatics, Drexel University

Zhao Li
Alibaba Group

ABSTRACT

Predicting the survival of cancer patients holds significant meaning for public health, and has attracted increasing attention in medical information communities. In this study, we propose a novel framework for cancer survival prediction named Multimodal Graph Neural Network (MGNN), which explores the features of real-world multimodal data such as gene expression, copy number alteration and clinical data in a unified framework. In order to explore the inherent relation, we first construct the bipartite graphs between patients and multimodal data. Subsequently, graph neural network is adopted to obtain the embedding of each patient on different bipartite graphs. Finally, a multimodal fusion neural layer is designed to fuse the features from different modal data. The output of our method is the classification of short term survival or long term survival for each patient. Experimental results on one breast cancer dataset demonstrate that MGNN outperforms all baselines. Furthermore, we test the trained model on lung cancer dataset, and the experimental results verify the strong robust by comparing with state-of-the-art methods.

KEYWORDS

Medical information retrieval, Cancer survival prediction, Graph neural networks, Multimodal

ACM Reference Format:
Jianliang Gao, Tengfei Lyu, Fan Xiong, Jianxin Wang, Weimao Ke, and Zhao Li. 2020. MGNN: A Multimodal Graph Neural Network for Predicting the Survival of Cancer Patients. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '20), July 25–30, 2020, Virtual Event, China.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3397271.3401214

1 INTRODUCTION

Medical Information Retrieval (IR) plays an important role in public health [1], and has attracted increasing attention in many applications such as cancer prognosis prediction and epidemiological prediction [2]. In medical IR area, cancer prognosis prediction has significant meaning since cancer has become one of the worldwide health problems. Furthermore, predicting the survival expectancy of cancer patients is a key problem among cancer prognosis prediction [3]. Cancer survival prediction can be formulated as a censored survival analysis problem, predicting both if and when an event (i.e. patient death) occurs within a given time period [4]. In medical IR area, the five-year survival rate is a common used index for cancer prognosis. For example, only slightly more than half of the newly diagnosed oral cancer patients will survive over five years [5].

Accurate prediction of the cancer survival could help clinicians make effective decisions and establish appropriate therapy programs. Meanwhile, it can spare a significant number of patients from receiving unnecessary treatment and its related expensive medical costs [6]. Prior works have tried to solve the problem of cancer prognosis prediction. Support vector machine-based recursive feature elimination approach is proposed for prognosis prediction by relying on gene expression data [7]. A recent interdisciplinary effort is to approach this problem from a deep learning perspective. A novel multimodal deep neural network by integrating multidimensional data for human breast cancer prognosis prediction is proposed [6]. Recently, deep learning has been an emerging methodology and provided a possible approach to improve the accuracy of cancer survival prediction [8]. However, there are two challenges to adopt deep learning for cancer survival prediction: (1) How to utilize the features of medical data from different modalities (e.g., gene expression and clinical). Most of prior works directly combine different types of data into deep learning model without utilizing the rich information of multimodal data. (2) How to utilize the structure information between patients and multimodal medical data. Most of previous works ignore the inherent relations between patients and medical data.

To cope with the above challenges, we design a Multimodal Graph Neural Network (MGNN) to predict the survival of patients. We utilize gene expression profile or the DNA copy number alteration (CNA) profile to construct a bipartite graphs, which explore the relation among them. Based on the bipartite graphs, a new graph neural network is proposed to obtain the embedding representation of each patient. Finally, the output is the classification of short term survival or long term survival based on a given threshold (usually adopt 5-year survival in the community). The main contributions of this paper are summarized as follows:

SIGIR '20, July 25–30, 2020, Virtual Event, China

ACM ISBN 978-1-4503-8016-4/20/07…$15.00
https://doi.org/10.1145/3397271.3401214

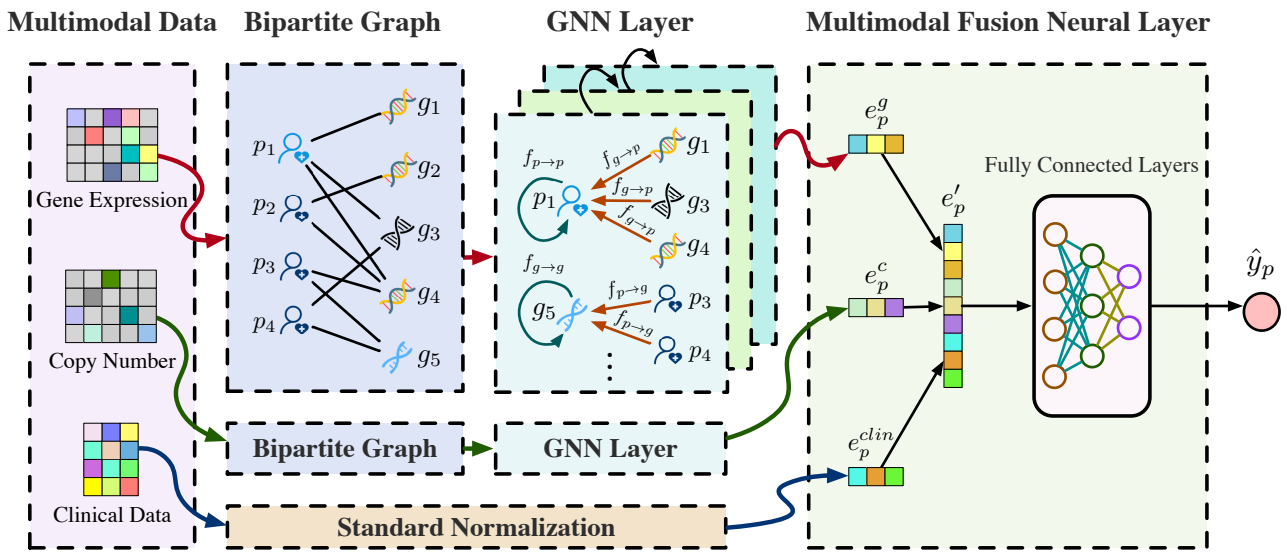

Figure 1: The overview of the proposed MGNN. MGNN is composed of three main parts: bipartite graph, GNN layer, multimodal fusion neural layer. With the input multimodal data such as gene expression and copy number, (1) bipartite graph further expresses their potential relation (e.g, the edges between patient p_i and gene expression g_j); (2) the GNN layer, which aggregates information of neighbors in bipartite graph by graph neural network model; and (3) the multimodal fusion neural layer, which fuses the features of multimodal data. For the output labels $\widehat{y}_p$, the short term survivors (less than 5 year survival) are labeled as 0 and long term survivors (more than 5 year survival) are labeled as 1.

- We highlight the critical importance of explicitly exploiting the multimodal data, and the inherent relation between patients and multimodal data.
- We propose a novel Multimodal Graph Neural Network (MGNN) for cancer survival prediction, which utilizes multimodal data in a unified framework. In the framework, the problem of cancer survival prediction is solved as a learning task of patient classification.
- We conduct experiments on real-world datasets. The results demonstrate the state-of-art performance of MGNN and its effectiveness and robustness for cancer survival prediction.

2 PROPOSED MGNN METHOD

In this section, we first define the problem formulation. Then, the detailed design of MGNN, which learns classification information from multimodal data, is shown in Figure 1. The overall structure is composed of three parts: the bipartite graph, the GNN layer, and the multimodal fusion neural layer.

2.1 Problem Formulation

In this study, multimodal data are composed of the gene expression profile data, the CNA profile data and the clinical data. It is expressed as follows:

$$\mathcal{X}_p = \{X_g, X_c, X_{clin}\} \in \mathbb{R}^{N\times(m+n+k)}, \tag{1}$$

where $X_g \in \mathbb{R}^{N\times m}$, $X_c \in \mathbb{R}^{N\times n}$ and $X_{clin} \in \mathbb{R}^{N\times k}$ stand for the gene expression profile data, the CNA profile data and the clinical data respectively. m, n and k represent the dimension of the gene data, the CNA data and the clinical data respectively and N is the number of patients. We first build bipartite graphs via $X_g \in \mathbb{R}^{N\times m}$ and $X_c \in \mathbb{R}^{N\times n}$, and then the node of bipartite graph aggregates information of neighbors by GNN model. Specifically, the final embedding of the patient can be shown as below:

$$\mathcal{E}_p = \{E_g, E_c, E_{clin}\} \in \mathbb{R}^{N\times(d_1+d_2+k)}, \tag{2}$$

where $E_g \in \mathbb{R}^{N\times d_1}$, $E_c \in \mathbb{R}^{N\times d_2}$ and $E_{clin} \in \mathbb{R}^{N\times k}$ respectively represent the patient embedding from multimodal data, d_1 and d_2 represent the size of the embedding dimension. $e'_p \in \mathcal{E}_p$ denotes the multimodal embedding vector for each patient.

The goal of the study is to distinguish short term survivors and long term survivors (usually adopt less 5-year survival), and the formal definition of the problem is as follows:

Input: multimodal data $\mathcal{X}_p = \{X_g, X_c, X_{clin}\}$.

Output: the predictive classification of short term survivors or long term survivors.

In order to predict the survival of cancer patients from multimodal data, we design a novel *Multimodal Graph Neural Network* (MGNN) as illustrated in Figure 1.

2.2 MGNN

In this subsection, we introduce the detailed design of the MGNN for predicting the survival of cancer patients.

Bipartite Graph. In order to establish the link between patients and multimodal, we utilize the gene expression profile and the copy number alteration (CNA) which belong to the multimodal data to construct bipartite graph. According to the previous method [9], the gene expression features are normalized and processed into three categories: under-expression (-1), over-expression (1), baseline

(0). For each patient, an edge will be built between the patient and the gene, only if the gene is not properly expressed (under-expression or over-expression). Finally we construct a patient-gene bipartite graph. Obviously, we can intuitively understand the gene expression affecting patients from the patient-gene bipartite graph.

For CNA data, we directly utilize the original data with five discrete values: homozygous deletion (-2); hemizygous deletion (-1); neutral/no change (0); gain (1); high level amplification (2). A patient-CNA bipartite graph is constructed by using the similar approach of the construction of patient-gene bipartite.

GNN Layer. After getting patient-gene bipartite graph, the initial representation matrix of patient-gene bipartite E is as follows:

$$E = [\underbrace{e_{p_1}^{(0)}, e_{p_2}^{(0)}, \cdots, e_{p_N}^{(0)}}_{patients\ embedding}, \underbrace{e_{g_1}^{(0)}, e_{g_2}^{(0)}, \cdots, e_{g_M}^{(0)}}_{genes\ embedding}], \tag{3}$$

where N and M stand for the number of patient and gene, respectively, $e_p^{(0)}, e_g^{(0)}$ are served as the initialization of patient embedding and gene embedding respectively.

In order to get the genetic information of patients more efficiently, we aggregate the messages propagated from the neighborhood of p to refine the embedding of p. Specifically, we define the aggregation function as below:

$$e_p^{(l+1)} = \sigma(f_{p\to p}(e_p^{(l)}) + \sum_{g\in\mathcal{N}_p} f_{g\to p}(e_p^{(l)}, e_g^{(l)})), \tag{4}$$

where $e_p^{(l+1)}$ denotes the embedding of patient p obtained at the $(l+1)^{th}$ GNN layer, $\mathcal{N}_p$ is neighbor set of patient p, $\sigma(\cdot)$ denotes activation function $LeakyReLU$, and $f(\cdot)$ is the representation encoding function that can be shown as below:

$$\begin{cases} f_{p\to p}(e_p^{(l)}) = W_1^{(l)} e_p^{(l)} \\ f_{g\to p}(e_p^{(l)}, e_g^{(l)}) = (W_1^{(l)} e_g^{(l)} + W_2^{(l)}(e_g^{(l)} \odot e_p^{(l)})) \end{cases} \tag{5}$$

where $W_1^{(l)}, W_2^{(l)} \in \mathbb{R}^{d_l \times d_{l+1}}$ are the trainable weight matrices, and $\odot$ denotes the element-wise product. $e_p^{(l)}$ and $e_g^{(l)}$ are the representation of patients and genes in l layer of GNN.

The embedding of patients based on the patient-gene bipartite graph, through the GNN layer, is e_p^g. Similarly, the embedding of patients e_p^c can be obtained by patient-CNA bipartite graph. The clinical data after standard normalization is denoted as e_p^{clin}.

Multimodal Fusion Neural Layer. After getting the representation e_p^g, e_p^c and e_p^{clin} for patient p, these embedding are linked together as the final multimodal embedding e_p'.

$$e_p' = e_p^g || e_p^c || e_p^{clin} \tag{6}$$

Then, the multimodal fusion embedding e_p' with multiple fully connected layers is used to predict the survival of cancer patients:

$$\hat{y}_p = softmax(\rho(W_3^{(l)} e_p'^{(l)} + b^{(l)})), \tag{7}$$

where $W_3^{(l)}$ is the trainable weight matrix, ρ denotes activation function $tanh$, $e_p'^{(l)}$ denotes the final multimodal representation for patient p at the layer l, and $b^{(l)}$ is the bias vector. Finally we use softmax function and obtain the final prediction score $\hat{y}_p$.

2.3 Optimization

For model optimization, our method could be trained with supervised setting. Besides, we use L2 regularization to prevent overfitting of our model. Based on the cross-entropy loss, the objective function could be defined as follows:

$$\begin{aligned} L(y_p, \hat{y}_p) = &-\frac{1}{N}\sum_{i=0}^{N}[y_{p_i} log\hat{y}_{p_i} - (1-y_{p_i}) log(1-\hat{y}_{p_i})] \\ &+ \frac{1}{\lambda}\sum_{l=1}^{L}\sum_{t=1}^{T}\sum_{i=1}^{d_l}\sum_{j=1}^{d_{l+1}} {w_t^l}_{ij}^{\,2}, \end{aligned} \tag{8}$$

where y_{p_i} is the actual label, $\hat{y}_{p_i}$ is the predictive scores and N is the batch size, L and T is the number of the embedding layers and trainable weight matrices W, respectively. $W_t^l = \{{w_t}_{ij}^l\}_{d_l \times d_{l+1}}$ is the l^{th} weight matrix.

3 EXPERIMENTS

In this section, we describe the experimental setups and results. The results obtained with our method are compared with four of the state-of-art methods for cancer survival prediction.

3.1 Experimental Setup

Datasets. We utilize well-established benchmark datasets[1], which are widely used for predicting the survival of breast cancer patients. The dataset is extracted from 1903 valid breast cancer patient's data and it contains multimodal data including gene expression profile, CNA profile and clinical data. In this work, the gene expression profile data include approximately 24369 genes and CNA profile data contain approximately 22544 genes in the breast cancer dataset.

Each patient has 35 clinical features, including age at diagnosis, lymph nodes examined positive, cancer type detailed, radio therapy etc. Finally, we utilize 28 clinical features in our experiment.

To comprehensively evaluate our proposed method, we adopt ten-fold cross validation in our experiments. Specifically, we randomly divide all patients into ten subsets. For each round of training, each subset will be used as a test set, and the remaining nine subsets are divided into a training set (80%) and a validation set (20%). The prediction score is the average of the output of ten rounds.

Parameter Settings. The proposed method is implemented with Tensorflow and optimized by Adam with a learning rate of 0.001. The parameters in Section 2 is set as $m = 64$, $n = 64$, $k = 28$, and $d_l = d_2 = 128$. The embedding $e_g^{(0)}$ and $e_c^{(0)}$ are 64 dimensional and initialized randomly. After obtaining the embedding e_p', it is filled into the full connection layers with 3 layers, and each hidden layer contains 200, 100 and 100 units.

Evaluation Metrics. We evaluate the MGNN with five metrics including the Area Under the Curve (AUC) of the receiver operating characteristics (ROC) curve, *Sensitivity* (*Sn*), *Accuracy* (*Acc*), *Precision* (*Pre*) and Matthew's correlation coefficient (*Mcc*).

3.2 Results and Analysis

We analyze the experimental results on breast cancer dataset to verify the effectiveness of MGNN, including performance comparison, robustness verification, and ablation tests.

[1]Datasets are available at https://www.cbioportal.org/

Table 1: Comparison of Acc, Pre, Sn, Mcc and AUC between LR, RF SVM, MDNNMD and MGNN methods (Train and test on breast cancer dataset)

Methods	Acc	Pre	Sn	Mcc	AUC
LR	0.760	0.549	0.183	0.209	0.663
RF	0.791	0.766	0.226	0.337	0.801
SVM	0.805	0.708	0.365	0.407	0.810
MDNNMD	0.826	0.749	0.450	0.486	0.845
MGNN	**0.940**	**0.953**	**0.969**	**0.837**	**0.970**

Performance Comparison. We compare the results of our method for breast cancer survival prediction with that of four state-of-the-art models including MDNNMD[6], SVM[7], RF[10] and LR[11]. We directly use the experimental results of MDNNMD. Table 1 show the experimental results of these models with the evaluation metrics including *Acc*, *Pre*, *Sn*, *Mcc* and *AUC* respectively. Compared to these baselines, it can be seen that the proposed MGNN achieves the best performance among all methods. For example, the accuracy of our method reaches 94%, which is higher 11.4% than the second high accuracy of *MDNNMD* method. It shows the power of MGNN method to extract multimodal data features. Besides, graph neural networks can effectively assemble bipartite graph structure interaction information for predicting the survival of cancer patients.

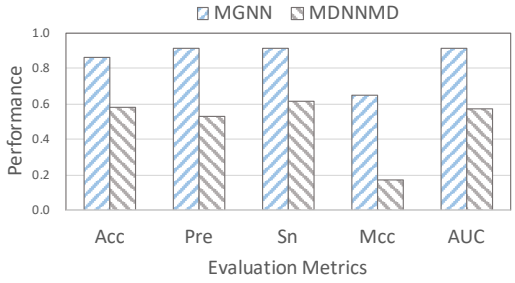

Figure 2: Robustness verification on lung cancer dataset.

Robustness Verification. To verify the robustness of the proposed MGNN, we apply our model to survival prediction for lung cancer patients. The format of the lung cancer dataset is the same as the breast cancer dataset. The lung cancer dataset contains CAN profile and clinical data of 917 lung cancer patients. As can be seen in Table 1, *MDNNMD* method achieves the best performance except our MGNN. Therefore, we select it as the baseline to compare the robustness. As shown in Figure 2, only CNA data is adopted to construct a bipartite graph, and MGNN is used to obtain the embedding representation of each patient on the bipartite graph. Compared to MDNNMD, our proposed method MGNN has best performance in terms of all evaluation metrics.

Ablation Test. To further explore the effect of multimodal fusion in our proposed model, we compare MGNN with GNN (gene) and GNN (CNA), where GNN (gene), GNN (CNA) are parts of our model and described as only using gene expression and copy number profile to make cancer survival prediction, respectively. As shown in Figure 3, all the evaluation metrics of MGNN achieves the best performance, which shows that MGNN can effectively fuse multimodal data from different dimensions.

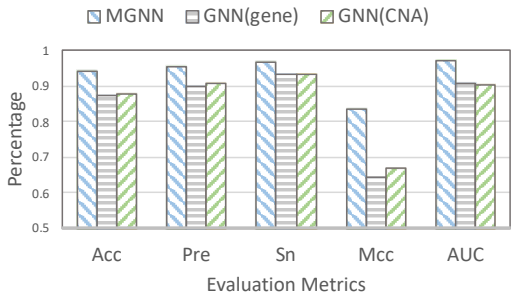

Figure 3: Ablation test results of the proposed MGNN.

4 CONCLUSIONS

In this paper, we propose a novel method MGNN for the cancer survival prediction of patients. The MGNN method utilizes the features of multimodal data with a unified framework. The experimental results showed the consistent performance improvement by the proposed method over the state-of-art baseline methods on real-world breast cancer dataset and lung cancer dataset.

ACKNOWLEDGMENTS

The work is supported by the National Natural Science Foundation of China under Grant No.: 61873288,61836016.

REFERENCES

[1] Xiaoli Wang, Rongzhen Wang, Zhifeng Bao, Jiayin Liang, and Wei Lu. Effective medical archives processing using knowledge graphs. In *SIGIR*, pages 1141–1144, 2019.
[2] Yuexin Wu, Yiming Yang, Hiroshi Nishiura, and Masaya Saitoh. Deep learning for epidemiological predictions. In *SIGIR*, pages 1085–1088, 2018.
[3] Yawen Xiao, Jun Wu, Zongli Lin, and Xiaodong Zhao. A deep learning-based multi-model ensemble method for cancer prediction. *Computer methods and programs in biomedicine*, 153:1–9, 2018.
[4] Anika Cheerla and Olivier Gevaert. Deep learning with multimodal representation for pancancer prognosis prediction. *Bioinformatics*, 35(14):i446–i454, 2019.
[5] Dong Wook Kim, Sanghoon Lee, Sunmo Kwon, and Woong Nam. Deep learning-based survival prediction of oral cancer patients. *Scientific Reports*, 9(1):1–10, 2019.
[6] Dongdong Sun, Minghui Wang, and Ao Li. A multimodal deep neural network for human breast cancer prognosis prediction by integrating multi-dimensional data. *TCBB*, 16(3):841–850, 2018.
[7] Xiaoyi Xu, Ya Zhang, Liang Zou, Minghui Wang, and Ao Li. A gene signature for breast cancer prognosis using support vector machine. In *International Conference on BioMedical Engineering and Informatics*, pages 928–931, 2012.
[8] Xiang Wang, Xiangnan He, Meng Wang, Fuli Feng, and Tat-Seng Chua. Neural graph collaborative filtering. In *SIGIR*, pages 165–174, 2019.
[9] PC Stone and S Lund. Predicting prognosis in patients with advanced cancer. *Annals of oncology*, 18(6):971–976, 2007.
[10] Cuong Nguyen, Yong Wang, and Ha Nam Nguyen. Random forest classifier combined with feature selection for breast cancer diagnosis and prognostic. *Journal of Biomedical Science and Engineering*, 6:551–560, 2013.
[11] Miles F Jefferson, Neil Pendleton, Sam B Lucas, and Michael A Horan. Comparison of a genetic algorithm neural network with logistic regression for predicting outcome after surgery for patients with nonsmall cell lung carcinoma. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, 79(7):1338–1342, 1997.