A Graph Attention Network Combining Radiomics and Clinical Data for Interpretable Prediction of Lymph Node Invasion in High-Grade Prostate Cancer

Maxence Larose[1,2,3] MAXENCE.LAROSE.1@ULAVAL.CA
[1] *Department of physics, engineering physics and optics, Université Laval, Québec (Qc), Canada*
[2] *CHU de Québec, Québec (QC), Canada*
[3] *Department of Computer Science, Université de Sherbrooke, Sherbrooke (Qc), Canada*
Nawar Touma[2] NAWAR.TOUMA.1@ULAVAL.CA
Nicolas Raymond[3] NICOLAS.RAYMOND2@USHERBROOKE.CA
Danahé LeBlanc[1,2] DANAHE.LEBLANC.1@ULAVAL.CA
Fatemeh Rasekh[2] FM.RASEKH@GMAIL.COM
Bertrand Neveu[2] BERTRAND.NEVEU@CRCHUDEQUEBEC.ULAVAL.CA
Hélène Hovington[2] HELENE.HOVINGTON@CRCHUDEQUEBEC.ULAVAL.CA
Martin Vallières[3] MARTIN.VALLIERES@USHERBROOKE.CA
Frédéric Pouliot[2] FREDERIC.POULIOT@FMED.ULAVAL.CA
Louis Archambault[1,2] LOUIS.ARCHAMBAULT@PHY.ULAVAL.CA

Abstract

This work proposes the use of a graph attention network (GAT) model combining radiomics and clinical data to improve the performance and interpretability of lymph node invasion (LNI) prediction in high-grade prostate cancer (PCa). Experiments were conducted using an in-house dataset of 170 high-grade PCa (Gleason $\geq$ 8), each with FDG-PET/CT images acquired prior to prostatectomy. To ensure interpretable connections between patients, the graph structure was constructed by merging the most important radiomic shape-based CT feature and first-order intensity-based PET feature to the clinically relevant features. The performance of the GAT model was compared to random forest (RF) and support vector machine (SVM) classifiers. On the 30 patients test set, the models reached {**AUC=0.765**, bACC= 0.705}, {AUC=0.748, bACC=0.66} and {AUC=0.725, **bACC=0.725**} for the GAT, RF and SVM models respectively. Even though SVM achieved higher balanced accuracy than GAT, the predictions made by the latter are more interpretable through the graph structure and attention mechanism.

Keywords: Prostate cancer, lymph node metastasis, GAT, multi-modal data, FDG-PET/CT, radiomics

1. Introduction

Prostate cancer (PCa) is the second most frequent cancer and the fifth leading cause of cancer death among men. To improve patient outcomes, PCa treatment must be personalized based on accurate prognosis. Nomograms that can identify patients at low lymph node invasion (LNI) risk according to preoperative information already exist (Cimino et al., 2017). However, a recent study shows that combining imaging and non-imaging data in a graph attention network (GAT) to predict various COVID-19 outcomes outperforms non-graph baselines in addition to providing insights into the decision-making process (Keicher et al., 2021). Based on this work, we developed a GAT model combining FDG-PET/CT images and clinical data to improve pre-treatment prognostic in PCa by accurately predicting LNI.

2. Methods

As detailed below, Figure 1 depicts the processing pipeline, the graph structure and the attention mechanism relating similar patients during inference.

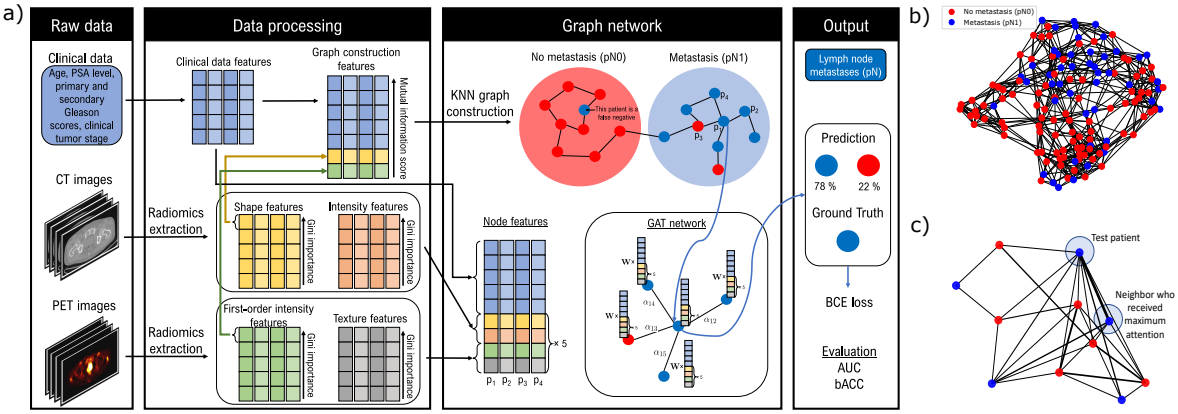

Figure 1: **(a)** GAT model that combines both radiomics features with clinical data to predict LNI. **(b)** Graph constructed by connecting each node to its 7 nearest neighbors. **(c)** Subgraph showing the attention mechanism for a single test patient.

Dataset: Patients in this cohort were all diagnosed with advanced prostate cancer with Gleason scores higher than 8 at biopsy. The clinical variables included age, pre-operative PSA score, primary and secondary Gleason scores and clinical stage. Each patient had pre-operative FDG-PET/CT images and the prostate was manually delineated on the CT.

Radiomic feature extraction and categories: A total of 195 radiomic markers were extracted from the images using `PyRadiomics`. The extracted radiomic features were separated into four categories based on modality-specific relevance and interpretability potential: (i) CT shape-based features; (ii) CT intensity-based features (first-order and texture); (iii) PET first-order intensity features; and (iv) PET texture features. The features of the four groups were independently ordered by Gini importance.

Graph structure construction: The graph structure was constructed using a k-nearest neighbor method based on mutual information task-weighted cosine similarity between patient features, with $k = 7$ neighbors for the final model (Figure 1b). To maximize the interpretability of the graph edges, only the clinical features combined with the single most important CT shape and PET first-order intensity features were used in this step.

Graph node features: To focus on features carrying useful information, the node features used for training the graph node prediction task were a combination of : (i) clinical features; and (ii) the five most important features from each of the four radiomics categories.

3. Experimental results

Experimental setup: Our proposed method is based on graph attention layers (GAT) (Veličković et al., 2018) and was implemented using the `Deep Graph Library`. The repre-

sentations of the nodes are updated by aggregating the 1-hop neighborhood information. The performance of the GAT model was compared to a random forest (RF) and support vector machine (SVM) classifiers. Class weights were adjusted during training to account for class imbalance (67% pN0, 33% pN1). The models were optimized using a stratified nested 5-fold cross-validation with 140 patients and evaluated on a test set of 30 patients.

Results: As shown in Table 1, the GAT model achieved a better AUC than the other two on the test set; however, DeLong test shows that the improvement is not significant ($\text{p-value}_{\text{GAT-RF}} = 0.85$, $\text{p-value}_{\text{GAT-SVM}} = 0.64$). The GAT also achieved a slightly lower balanced accuracy (bACC) than the SVM. Nonetheless, the predictions made by GAT have the advantage of being more interpretable by identifying the neighbors that receive maximal attention, as shown in Figure 1c, so the impact it could have in a clinical setting is enhanced.

Table 1: Results on validation and test sets.

Dataset	Models	AUC	bACC	Sensitivity	Specificity
Validation	GAT	**0.68±0.09**	0.65±0.07	0.59±0.06	0.64±0.10
	RF	0.66±0.04	0.61±0.06	0.32±0.11	0.91±0.07
	SVM	**0.68±0.07**	**0.68±0.07**	0.60±0.13	0.76±0.07
Test	GAT	**0.765**	0.705	0.68	0.73
	RF	0.748	0.660	0.50	0.82
	SVM	0.725	**0.725**	0.70	0.75

4. Conclusion

We developed a GAT model that uses the combination of imaging and non-imaging data to predict LNI in high-grade PCa. We showed that the model performs well and provides a natural way to interpret the predictions. For future work, automatic prostate segmentation would facilitate the use of such a model in a clinical setting. In addition, an advanced search of model hyperparameters and larger dataset could improve prediction performance.

Acknowledgments

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and by the Fonds de Recherche du Québec (FRQNT).

References

Sebastiano Cimino, Giulio Reale, Tommaso Castelli, et al. Comparison between briganti, partin and mskcc tools in predicting positive lymph nodes in prostate cancer: a systematic review and meta-analysis. *Scandinavian Journal of Urology*, 51(5):345–350, 2017.

Matthias Keicher, Hendrik Burwinkel, David Bani-Harouni, et al. U-GAT: multimodal graph attention network for COVID-19 outcome prediction. *CoRR*, abs/2108.00860, 2021.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, et al. Graph attention networks. In *International Conference on Learning Representations*, 2018.