

KOAL: Knowledge-Driven Prostate Cancer Grading with Ordinal-Aware Learning

Zheng Guo^{1†}, Jiaqi Cui^{1†}, Haocheng Xiong¹, Jize Han², Bo Liu¹, Qianwen Zhang^{3(✉)}, Rui Chen^{4(✉)}, and Yan Wang^{1(✉)}

¹ School of Computer Science, Sichuan University, China
wangyanscu@hotmail.com

² China Mobile (Suzhou) Software Technology Company Limited, China

³ Shanghai Changhai Hospital, Naval Medical University, China

⁴ Renji Hospital, Shanghai Jiao Tong University School of Medicine, China
drchenrui@foxmail.com

Abstract. Lesion-level prediction of Gleason Grade Group (GGG) from multiparametric MRI (mpMRI) can support prostate cancer risk assessment and reduce unnecessary invasive biopsies. Existing GGG prediction methods face two major limitations. First, they often overlook non-image information critical for GGG prediction, including age, prostate-specific antigen (PSA), and expert priors embedded in radiology reports. Second, they tend to oversimplify GGG as flat categorical labels, failing to account for its intrinsic hierarchy of primary and secondary Gleason patterns. To this end, we propose a novel **Knowledge-Driven Ordinal-Aware Learning (KOAL)** framework with three synergistic modules. Specifically, the Clinical-Context Modulation (CCM) module uses clinical variables (e.g., age and PSA) to dynamically modulate discriminative image representations. The Knowledge-Guided Prototype Alignment (KGPA) module leverages an LLM to extract group-specific expert knowledge from training radiology reports and clinical guidelines, producing offline semantic anchors describing grade-specific radiological findings without requiring patient-specific reports at inference. Through prototype contrastive alignment, patient-specific mpMRI representations are matched with these anchors to promote pathology-aligned representation learning. The Hierarchical Ordinal-aware Constraints (HOC) module decouples primary and secondary Gleason pattern prediction and maps their probabilistic outputs to GGG via a Differentiable Bio-logic Mapping Layer (DBML), ensuring pathological grading consistency. Experiments on public PI-CAI and in-house datasets demonstrate that KOAL outperforms state-of-the-art methods. Code is available at: <https://github.com/Gother-GZ/KOAL>.

Keywords: Prostate Cancer; Gleason Grading Prediction; Multi-modal Learning; Ordinal Regression; Knowledge-Driven Learning.

1 Introduction

Prostate cancer (PCa) is one of the most prevalent malignancies in men worldwide [1]. In clinical practice, tumor aggressiveness is central to diagnosis and treatment, and is

[†] These authors contributed equally.

usually evaluated using the Gleason grading system [2]. This system assigns grades to the dominant (i.e., primary) and the next-most frequent (i.e., secondary) patterns, which together form the Gleason score (e.g., 3+3 to 5+5) [2]. Given the established association between Gleason score and patient prognosis, the International Society of Urological Pathology (ISUP) further introduced a simplified grading system, i.e., the Gleason Grade Group (GGG, ranging from 1 to 5) [3]. However, as a pathological gold standard, obtaining GGG typically necessitates invasive biopsy or surgical resection, which is associated with procedure-related complications [4]. Moreover, inconsistencies often exist between biopsy-based GGG and whole-mount pathological analysis [5]. Therefore, reliable non-invasive GGG prediction is highly desirable for supporting prostate cancer risk stratification and reducing unnecessary biopsies.

In clinical practice, multiparametric magnetic resonance imaging (mpMRI), including T2-weighted imaging (T2), diffusion-weighted imaging (DWI), and apparent diffusion coefficient (ADC), has been widely adopted for non-invasive prostate cancer assessment [6,7]. Nevertheless, accurate prediction of histopathological GGG from mpMRI remains challenging due to intra-tumor heterogeneity, inter-observer variability, and subtle imaging-pathology correlations [8]. Early studies employed radiomics-based approaches [9,10,11], which extracted handcrafted features from mpMRI to establish a mapping between imaging phenotypes and GGG. However, these methods often require precise lesion segmentation and expert-defined feature engineering, which are labor-intensive and exhibit limited generalization capabilities across different centers. More recently, deep learning (DL) techniques have been explored to improve GGG prediction and prostate cancer analysis [12,13,14,15,16]. In addition, given the inherent ordinal nature of GGG (1-5), several studies have explored modeling ordinal relationships among grade groups. In this vein, Xu *et al.* [17] proposed a Poisson ordinal network for Gleason group estimation on MRI to improve consistency and robustness.

Despite these improvements, existing methods still face two key limitations. First, these approaches primarily rely on image-only mpMRI inputs for GGG prediction. This paradigm, however, often overlooks expert knowledge conveyed in radiology reports, which reflect radiologists' case-by-case assessments that are difficult to infer from images alone. In addition, clinical variables such as age and prostate-specific antigen (PSA), which are well-established risk indicators for prostate cancer [21], are often not incorporated, potentially limiting prediction accuracy. Second, existing methods either model GGG as flat categorical labels or, at best, as a single ordinal variable. Although these formulations recognize the ordinal nature of GGG, they treat it as an indivisible entity and fail to account for its intrinsic compositional structure, defined by primary and secondary Gleason patterns. As both patterns are ordinal, ignoring this fine-grained pattern-level structure results in an overly coarse modeling of disease severity.

Motivated to tackle these limitations, we propose a **Knowledge-Driven Ordinal-Aware Learning (KOAL)** framework for lesion-level GGG prediction in prostate cancer after suspicious-lesion localization. KOAL consists of three synergistic components, i.e., a Clinical-Context Modulation (CCM) module, a Knowledge-Guided Prototype Alignment (KGPA) module, and a Hierarchical Ordinal-aware Constraints (HOC) module. Specifically, the CCM module uses clinical variables (e.g., PSA and age) as patient-specific context to condition a cross-attention mechanism, dynamically

modulating image representations. Moreover, the KGPA module leverages a large language model (LLM) to extract grade group-specific knowledge from training radiology reports and guidelines. This process constructs a group-level knowledge bank that captures radiologists’ population-level priors. From these priors, inference-ready semantic anchors are derived to describe the radiological findings of each grade group. These anchors are constructed offline and reused during inference, so patient-specific radiology reports are not required at test time. Through prototype contrastive alignment, patient-specific mpMRI representations are aligned with these semantic anchors, promoting pathology-aligned representation learning. In addition, to model the fine-grained ordinal structure of GGG, HOC decouples GGG prediction into primary and secondary Gleason pattern prediction, and maps their probabilistic outputs to the final GGG via a Differentiable Bio-logic Mapping Layer (DBML) based on ISUP-defined combinatorial rules, thereby enforcing consistency with clinical grading logic. Through these designs, KOAL effectively exploits the potential of heterogeneous imaging (i.e., mpMRI) and non-imaging information (i.e., radiology reports and clinical variables) with fine-grained ordinal supervision for improved GGG prediction.

The main contributions of this paper are summarized as follows. (1) We propose KOAL, the first knowledge-driven ordinal-aware learning framework that integrates heterogeneous information from mpMRI, radiology reports, and clinical variables for GGG prediction in prostate cancer. (2) The KGPA module leverages an LLM to extract grade-group knowledge from training reports and guidelines, deriving semantic anchors to align patient-specific mpMRI representations with population-level radiological priors. (3) The HOC module decouples primary and secondary Gleason patterns, mapping them to the final GGG using the DBML, thereby enforcing compositional ordinal constraints consistent with clinical grading. (4) Extensive experiments on the public PI-CAI dataset and an in-house dataset demonstrate that KOAL outperforms existing state-of-the-art approaches.

2 Methodology

The KOAL framework is illustrated in Fig. 1. It leverages three core modules, i.e., the Clinical-Context Modulation (CCM) module, the Knowledge-Guided Prototype Alignment (KGPA) module, and the Hierarchical Ordinal-aware Constraints (HOC) module, to integrate mpMRI, clinical variables, and report-derived semantic priors for lesion-level GGG prediction after suspicious-lesion localization. The details of each module are described in the following subsections.

2.1 Clinical Context Modulation (CCM)

In clinical practice, clinicians consider both mpMRI and clinical variables when assessing tumor aggressiveness and estimating Gleason grade. However, existing methods struggle to integrate mpMRI modalities with distinct focuses (i.e., anatomical structures from T2 and functional characteristics from DWI/ADC) and to fuse imaging with clinical variables. This results in suboptimal feature representations and limited

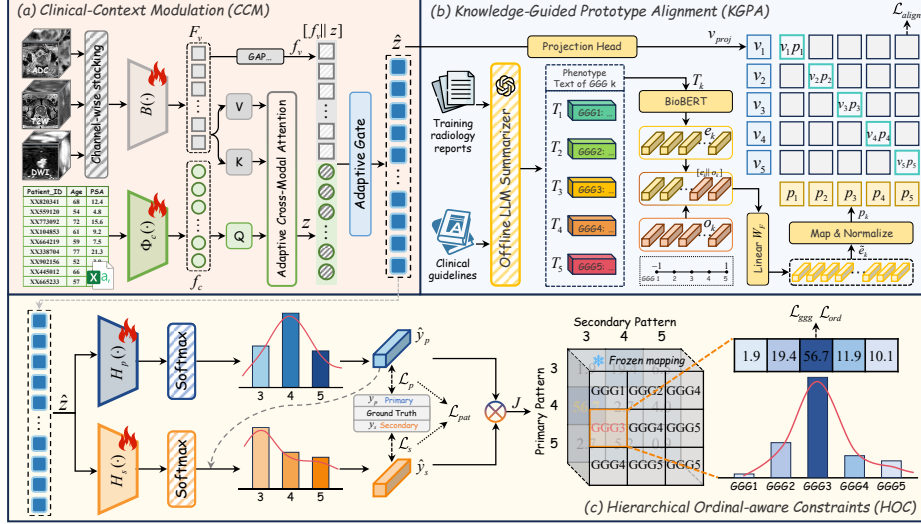


Fig. 1. Overview of the proposed KOAL framework. The architecture consists of three synergistic modules, including (a) Clinical-Context Modulation (CCM); (b) Knowledge-Guided Prototype Alignment (KGPA); and (c) Hierarchical Ordinal-aware Constraints (HOC).

prediction accuracy. As shown in Fig. 1(a), the CCM module tackles this challenge by combining channel-wise early fusion with clinically guided feature recalibration.

mpMRI Image Encoding. To exploit the complementarity of mpMRI, we adopt a channel-wise early fusion strategy. Let $\mathcal{S} = \{I_m \in \mathbb{R}^{D \times H \times W}\}_{m \in M}$ denote a set of spatially aligned mpMRI volumes, where $M = \{T2, ADC, DWI\}$. First, we stack these volumes along the channel dimension to form $X \in \mathbb{R}^{3 \times D \times H \times W}$. Then, a 3D ResNet-50 backbone $\mathcal{B}(\cdot)$ [22] extracts visual feature maps $F_v = \mathcal{B}(X)$. We further apply Global Average Pooling (GAP) and a projection head to obtain a pooled visual embedding $f_v = \mathcal{P}_{map}(\text{GAP}(F_v)) \in \mathbb{R}^d$, where $\mathcal{P}_{map}(\cdot)$ denotes an additional mapping layer for reduction and normalization operations.

Clinical Variables Encoding. Given a set of clinical variables $c \in \mathbb{R}^L$, where L is the length of the set, we employ a non-linear semantic mapping network $\Phi_c(\cdot)$ to transform c into a context-rich semantic embedding f_c . This transformation is achieved through layer normalization $\text{LN}(\cdot)$ and multi-layer perceptron projection $\mathcal{P}_{mlp}(\cdot)$, formulated as $f_c = \Phi_c(c) = \mathcal{P}_{mlp}(\delta(\text{LN}(c))) \in \mathbb{R}^d$, where δ denotes the GELU activation function. This mapping projects f_c into a shared semantic space with the image feature f_v for subsequent cross-modal interaction.

Adaptive Cross-Modal Attention (ACMA). To refine image representations using clinical variables, we develop the ACMA mechanism. Concretely, we treat f_c as the query $Q_c \in \mathbb{R}^{1 \times d}$, and flatten F_v into N token features to form the key $K_v \in \mathbb{R}^{N \times d}$ and value $V_v \in \mathbb{R}^{N \times d}$. The clinical-guided context z is computed as:

$$z = \mathcal{A}(f_c, F_v) = \text{Softmax}(Q_c K_v^\top / \sqrt{d}) V_v, \quad (1)$$

where the Softmax operation is applied over the N tokens. To avoid over-reliance on clinical variables, we further introduce an adaptive gating unit to balance the original and clinical-modulated features, as follows:

$$\hat{\mathbf{z}} = \text{LN}(G \odot \mathbf{z} + (1 - G) \odot \mathbf{f}_v), \quad G = \sigma(\mathbf{W}_G[\mathbf{f}_v \parallel \mathbf{z}]), \quad (2)$$

where $G \in \mathbb{R}^d$ is a channel-wise gating vector and $\mathbf{W}_G \in \mathbb{R}^{d \times 2d}$ is a learnable matrix. $\odot$ denotes element-wise multiplication, $[\cdot \parallel \cdot]$ represents concatenation, and $\sigma(\cdot)$ is the sigmoid function. The resulting $\hat{\mathbf{z}}$ encodes anatomical and functional information from mpMRI while being modulated by clinical risk factors such as PSA and age, thereby serving as a robust multimodal basis for subsequent knowledge learning and GGG prediction.

2.2 Knowledge-Guided Prototype Alignment (KGPA)

Although the recalibrated representation $\hat{\mathbf{z}}$ incorporates multimodal cues, optimizing it only with categorical supervision still treats the GGG grades as flat and independent classes, overlooking their intrinsic severity order ($\text{GGG} \in \{1, 2, 3, 4, 5\}$). As a result, the learned embeddings may become scattered and overlap across adjacent grades. To address this, KGPA leverages training radiology reports to construct grade-specific semantic anchors $\mathcal{P} = \{\mathbf{p}_k\}_{k=1}^K$ with $K = 5$. These anchors act as report-driven prototypes to regularize the latent feature space and promote ordinal-consistent representations.

LLM-based Expert Knowledge Extraction. To incorporate expert knowledge from training radiology reports and clinical guidelines, we use an LLM as an offline knowledge extractor. Specifically, we prompt the LLM to aggregate and summarize recurrent radiological findings and guideline-based grading priors for each grade group, producing a standardized population-level phenotype description T_k for the k -th grade group. To improve clinical validity and consistency, established clinical guidelines are incorporated into the prompt as structural constraints (e.g., a unified template specifying required fields and terminology). The resulting descriptions are converted into semantic anchors and reused at inference without requiring patient-specific reports.

Ordinal-Aware Learnable Prototypes. To bridge report-derived general knowledge with patient-specific mpMRI, we design a learnable prototype bank $\mathbf{E} = \{\mathbf{e}_k \in \mathbb{R}^{d_{\text{emb}}}\}_{k=1}^K$, where each prototype $\mathbf{e}_k$ is initialized from the embedding of T_k and is subsequently refined during training. To encode the natural progression of disease severity in GGG (from 1 to 5), we assign each grade an ordinal positional embedding $\mathbf{o}_k \in \mathbb{R}^{d_{\text{ord}}}$, within the range $[-1, 1]$. The ordinal-aware representation $\tilde{\mathbf{e}}_k$ is computed as:

$$\tilde{\mathbf{e}}_k = \mathbf{W}_F[\mathbf{e}_k \parallel \mathbf{o}_k] \in \mathbb{R}^d, \quad \mathbf{p}_k = \Phi_{\text{pro}}(\tilde{\mathbf{e}}_k) / \|\Phi_{\text{pro}}(\tilde{\mathbf{e}}_k)\|_2 \in \mathbb{R}^d, \quad (3)$$

where $\tilde{\mathbf{e}}_k$ is the fused prototype representation, $\mathbf{W}_F$ is a learnable linear projection layer, and $\Phi_{\text{pro}}(\cdot)$ is a mapping network. As training progresses, these anchors are refined via the following prototype contrastive alignment and gradually evolve into grade-specific centers in the embedding space.

Prototype Contrastive Alignment. To align patient-specific features $\hat{\mathbf{z}}$ from CCM with the knowledge anchors, we adopt the prototype contrastive loss $\mathcal{L}_{\text{align}}$ as follows:

$$\mathcal{L}_{\text{align}} = -\log \frac{\exp(\langle \mathbf{v}_{\text{proj}}, \mathbf{p}_y \rangle / \tau)}{\sum_{k=1}^K \exp(\langle \mathbf{v}_{\text{proj}}, \mathbf{p}_k \rangle / \tau)}, \quad (4)$$

where $\mathbf{v}_{\text{proj}}$ represents the projected features of $\hat{\mathbf{z}}$, $y \in \{1, \dots, K\}$ is the ground-truth label (thus $\mathbf{p}_y$ denotes the corresponding positive anchor). This constraint encourages an

ordinal topology in the feature space, faithfully mirroring the progression of cancer severity, ranging from mild (GGG=1) to advanced pathological stages (GGG=5).

2.3 Hierarchical Ordinal-aware Constraints (HOC)

Pattern Decoupling and Prediction. The GGG is defined by the combination of primary and secondary Gleason patterns according to ISUP standards. Inspired by this, the HOC module transforms conventional flat classification of GGG into a hierarchical learning process. To achieve this, we first decouple GGG prediction into primary and secondary pattern prediction using a dual-stream network. Given $\hat{\mathbf{z}}$ from CCM, the dual-stream network employs two independent pattern prediction layers, denoted as $H_p(\cdot)$ and $H_s(\cdot)$, to predict the probability distributions of the primary pattern $\hat{y}_p$ and the secondary pattern $\hat{y}_s$, respectively, as follows:

$$\hat{y}_p = \text{Softmax}(H_p(\hat{\mathbf{z}})), \quad \hat{y}_s = \text{Softmax}(H_s(\hat{\mathbf{z}})), \quad (5)$$

This process encourages the model to learn discriminative and fine-grained visual cues related to the dominant and subordinate pathological patterns, rather than simply memorizing global labels.

Differentiable Bio-logic Mapping (DBML). DBML maps the decoupled predictions to the final GGG distribution. We first compute the joint probability matrix $\mathbf{J} \in \mathbb{R}^{3 \times 3}$ as $\mathbf{J} = \hat{y}_p \otimes \hat{y}_s$, where $\otimes$ signifies the outer product. Each element $\mathbf{J}_{ij}$ corresponds to the joint probability of the primary pattern i and the secondary pattern j (where $i, j \in \{3, 4, 5\}$). Subsequently, we calculate the GGG distribution as $\hat{y}_{ggg} = \text{vec}(\mathbf{J}) \cdot \mathbf{W}_m$, where $\mathbf{W}_m$ is a frozen sparse matrix encoding the ISUP mapping logic. This differentiable operation enables end-to-end structured learning by propagating gradients from the final GGG supervision back to the decoupled pattern predictors in the dual-stream network, thus accurately reflecting the clinical grading process.

Hierarchical Ordinal Loss. To supervise the ordinal and hierarchical prediction of GGG, we develop the hierarchical ordinal loss, which consists of a pattern-level loss $\mathcal{L}_{pat}$, a grade-level loss $\mathcal{L}_{ggg}$, and a composite ordinal loss $\mathcal{L}_{ord}$. Specifically, the pattern-level loss aims to provide fine-grained supervision for the decoupled primary and secondary Gleason pattern predictions, formulated as:

$$\mathcal{L}_{pat} = CE(\hat{y}_p, y_p) + CE(\hat{y}_s, y_s), \quad (6)$$

where $CE(\cdot)$ represents the cross-entropy loss.

The grade-level loss applies an ordinal-sensitive regression loss [23] to supervise the final GGG prediction as:

$$\mathcal{L}_{ggg} = -\frac{1}{K-1} \sum_{j=1}^{K-1} [\mathbb{I}(y > j) \log \sigma(f_j) + (1 - \mathbb{I}(y > j)) \log (1 - \sigma(f_j))], \quad (7)$$

where f_j is the j -th ordinal logit, K is the number of Grade Groups, $y \in \{1, \dots, K\}$ is the ground-truth GGG label, and $\sigma(\cdot)$ is the sigmoid function.

The composite ordinal loss is adopted to adaptively penalize rank violations as:

$$\mathcal{L}_{ord} = \frac{1}{(K-1)^2} \sum_{k=1}^K \hat{y}_{ggg}[k] \cdot (k - y)^2, \quad (8)$$

where $\hat{y}_{ggg}[k]$ denotes the predicted probability on grade k . Notably, the quadratic term $(k - y)^2$ makes the penalty distance-aware, assigning larger costs to farther rank errors (e.g., $|k - y| = 4$ incurs 16 times the penalty of $|k - y| = 1$).

Finally, the total objective is the weighted sum of the above losses, formulated as:

$$\mathcal{L}_{total} = \mathcal{L}_{pat} + \lambda_{ggg}\mathcal{L}_{ggg} + \lambda_{ord}\mathcal{L}_{ord} + \lambda_{align}\mathcal{L}_{align}, \quad (9)$$

where λ_{ggg} , λ_{ord} , and λ_{align} are the hyperparameters to balance these terms.

3 Experiments and Results

3.1 Experimental Settings

Datasets. We trained and evaluated KOAL on the public PI-CAI dataset [24] of 643 patients and an in-house dataset of 1002 patients. For each lesion-centered sample, we collected mpMRI sequences, crucial clinical variables, and ground truth GGG labels with primary and secondary Gleason patterns. All MRI sequences were resampled to $20 \times 320 \times 320$ and cropped into $20 \times 96 \times 96$ lesion-centered ROI volumes to reduce background redundancy. All experiments used a fixed patient-level 80%/20% train/test split to avoid patient overlap.

Evaluation Metrics. We employ Quadratic Weighted Kappa (QWK) as the primary evaluation metric. We also report overall Accuracy (ACC) and macro-AUC to evaluate discriminative capability across different grades.

Implementation Details. All experiments were implemented in PyTorch on a single RTX 4090 GPU (24GB). The model was trained for 200 epochs (batch size 16) using the AdamW optimizer with an initial learning rate of $1e-4$. The hyperparameters λ_{ggg} , λ_{ord} , and λ_{align} in Eq. (9) are empirically set to 1.0, 0.5, and 0.1, respectively.

3.2 Comparison with State-of-the-Art Methods

To validate the effectiveness of KOAL, we compare it with (1) image-only methods, including PON [17], CORAL [23], AOR-DR [25], and SupCon [26]; and (2) multi-modal methods, including HyperFusion [28], PPAL [20], DSPA-MIL [27], XCoOp

Table 1. Quantitative comparison with nine state-of-the-art methods on the public PI-CAI dataset and the private In-House dataset. I: Imaging modality (mpMRI); T: Tabular clinical variables (e.g., PSA, age). The best results are highlighted in **bold**, and the second-best are underlined.

Methods	M		PI-CAI			In-House		
	I	T	QWK(%)	AUC(%)	ACC(%)	QWK(%)	AUC(%)	ACC(%)
PON [17]	✓		65.42	76.62	47.42	62.60	68.31	51.66
CORAL [23]	✓		77.06	84.18	56.70	60.11	73.23	42.38
AOR-DR [25]	✓		74.05	81.81	57.14	65.80	80.57	56.70
SupCon [26]	✓		79.70	89.42	72.16	67.20	83.92	64.12
HyperFusion [28]	✓	✓	82.66	83.82	60.82	77.96	62.96	59.60
PPAL [20]	✓	✓	82.64	92.07	69.41	64.62	82.44	57.25
DSPA-MIL [27]	✓	✓	76.19	85.61	70.10	53.49	76.70	52.98
XCoOp [19]	✓	✓	78.78	88.90	71.13	75.93	86.18	67.91
DAFT [18]	✓	✓	83.28	91.32	75.26	79.17	89.42	70.99
KOAL (Ours)	✓	✓	88.99	<u>91.93</u>	81.63	88.19	<u>89.38</u>	75.38

[19], and DAFT [18]. All baselines use the same split, mpMRI input, GGG target, and metrics. The results are presented in Table 1 and Fig. 2.

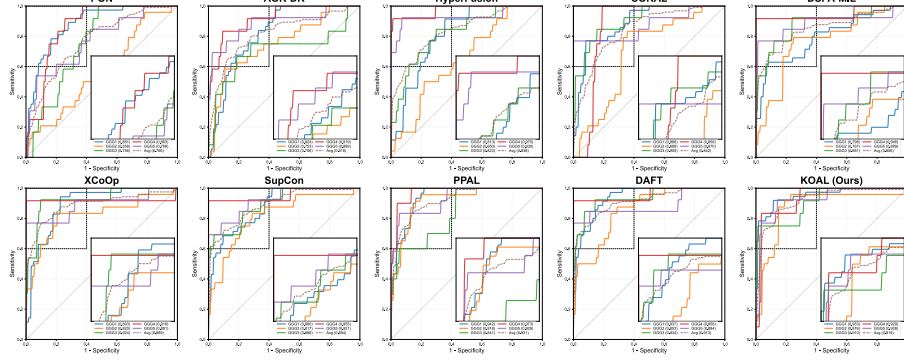


Fig. 2. Comparison of ROC curves between KOAL and competing methods on PI-CAI dataset.

Quantitative Results. On PI-CAI, KOAL achieves the highest QWK (88.99%) and ACC (81.63%), alongside a 91.93% AUC. DAFT ranks second in QWK/ACC, while PPAL obtains a marginally higher AUC. To evaluate robustness against clinical heterogeneity, we further validate on an In-House dataset with greater inter-patient variability. While baseline performance declines substantially on this dataset (e.g., PPAL’s QWK decreases from 82.64% to 64.62%), KOAL maintains robust performance, yielding 88.19% QWK, 75.38% ACC, and 89.38% AUC. These results underscore KOAL’s superior generalizability for clinical GGG prediction across varying data distributions.

3.3 Ablation Study

We conduct ablation studies on PI-CAI and the In-House datasets. As shown in Table 2, the mpMRI-only baseline performs the worst. The CCM module yields substantial gains ($\uparrow 10.27\%$ QWK on PI-CAI; $\uparrow 12.82\%$ on In-House). The KGPA module further boosts performance by using offline report-derived semantic anchors to align patient-specific mpMRI representations with grade-level priors. While standard ordinal loss (without primary/secondary decoupling) highlights the benefit of ordinal supervision, integrating the HOC module achieves optimal results, confirming the advantage of explicitly modeling the intrinsic compositional order of primary and secondary Gleason patterns in GGG prediction.

Table 2. Ablation study on the PI-CAI dataset and the In-House dataset. M: MpMRI (Baseline); C: CCM module; K: KGPA module; H: HOC module; O: standard ordinal loss treating GGG as a single ordered label (without primary/secondary decoupling).

Variants					PI-CAI			In-House		
M	C	K	O	H	QWK(%)	AUC(%)	ACC(%)	QWK(%)	AUC(%)	ACC(%)
✓					65.51	77.09	39.80	56.07	72.51	42.38
✓	✓				75.78	78.78	44.90	68.89	82.06	58.28
✓	✓	✓			83.67	88.29	68.37	74.17	83.46	57.62
✓	✓	✓	✓		85.04	89.83	74.23	84.60	84.74	68.90
✓	✓	✓		✓	88.99	91.93	81.63	88.19	89.38	75.38

4 Conclusion

In this paper, we propose KOAL, a novel knowledge-driven and ordinal-aware framework for non-invasive GGG prediction using mpMRI, clinical variables, and report-derived priors. Specifically, we introduce the CCM module, which modulates mpMRI representations based on patient-specific clinical variables. Meanwhile, the KGPA module uses offline semantic anchors derived from training reports and clinical guidelines to align image representations with grade-level priors. Finally, the HOC module models the hierarchical structure of GGG via pattern decoupling and DBML mapping. Experimental results on two prostate cancer datasets validate that our method achieves state-of-the-art performance, suggesting that KOAL is a promising tool for non-invasive prostate cancer assessment.

Acknowledgments. This work is supported by National Natural Science Foundation of China (NSFC 62371325, 82272905), Sichuan Science and Technology Program 2025NSFJQ0050, Sichuan Science and Technology Program 2024ZDZX0018.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Sung, H., Ferlay, J., Siegel, R.L., et al.: Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 71(3), 209–249 (2021).
2. Gleason, D.F., Mellinger, G.T.: Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging. *J. Urol.* 111(1), 58–64 (1974).
3. Epstein, J.I., Egevad, L., Amin, M.B., et al.: The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason grading of prostatic carcinoma: Definition of grading patterns and proposal for a new grading system. *Am. J. Surg. Pathol.* 40(2), 244–252 (2016).
4. Loeb, S., Vellekoop, A., Ahmed, H.U., et al.: Systematic review of complications of prostate biopsy. *Eur. Urol.* 64(6), 876–892 (2013).
5. Epstein, J.I., Feng, Z., Trock, B.J., et al.: Upgrading and downgrading of prostate cancer from biopsy to radical prostatectomy: incidence and predictive factors using the modified Gleason grading system and factoring in tertiary grades. *Eur. Urol.* 61(5), 1019–1024 (2012).
6. Cornford, P., van den Bergh, R.C.N., Briers, E., et al.: EAU-EANM-ESTRO-ESUR-ISUP-SIOG guidelines on prostate cancer—2024 update. Part I: Screening, diagnosis, and local treatment with curative intent. *Eur. Urol.* 86(2), 148–163 (2024).
7. Ahmed, H.U., El-Shater Bosaily, A., Brown, L.C., et al.: Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): A paired validating confirmatory study. *Lancet* 389(10071), 815–822 (2017).
8. Li, Y., Daho, M.E.H., Conze, P.H., et al.: A review of deep learning-based information fusion techniques for multimodal medical image classification. *Comput. Biol. Med.* 177, 108635 (2024).

9. Fehr, D., Veeraraghavan, H., Wibmer, A., et al.: Automatic classification of prostate cancer Gleason scores from multiparametric magnetic resonance images. *Proc. Natl. Acad. Sci. U. S. A.* 112(46), E6265–E6273 (2015).
10. Chaddad, A., Niazi, T., Probst, S., et al.: Predicting Gleason score of prostate cancer patients using radiomic analysis. *Front. Oncol.* 8, 630 (2018).
11. Lomer, N.B., Ashoobi, M.A., Ahmadzadeh, A.M., et al.: MRI-based radiomics for predicting prostate cancer grade groups: A systematic review and meta-analysis of diagnostic test accuracy studies. *Acad. Radiol.* 32(6), 3429–3452 (2025).
12. Cao, R., Mohammadian Bajgiran, A., Afshari Mirak, S., et al.: Joint prostate cancer detection and Gleason score prediction in mp-MRI via FocalNet. *IEEE Trans. Med. Imaging* 38(11), 2496–2506 (2019).
13. Santhirasekaram, A., Winkler, M., Rockall, A., et al.: Robust prostate disease classification using transformers with discrete representations. *Int. J. Comput. Assist. Radiol. Surg.* 20(1), 11–20 (2025).
14. Zhao, L., Bao, J., Wang, X., et al.: Detecting adverse pathology of prostate cancer with a deep learning approach based on a 3D Swin-Transformer model and biparametric MRI: A multicenter retrospective study. *J. Magn. Reson. Imaging* 59(6), 2101–2112 (2024).
15. Shao, L., Liu, Z., Liu, J., et al.: Patient-level grading prediction of prostate cancer from mp-MRI via GMINet. *Comput. Biol. Med.* 150, 106168 (2022).
16. Wilson, P.F.R., To, M.N.N., Jamzad, A., et al.: ProstNFound: Integrating foundation models with ultrasound domain knowledge and clinical context for robust prostate cancer detection. In: *MICCAI 2024. LNCS*, vol. 15006, pp. 499–509. Springer, Cham (2024).
17. Xu, Y., Wang, Y., Shen, Z., et al.: Poisson ordinal network for Gleason group estimation using bi-parametric MRI. In: Linguraru, M.G., et al. (eds.) *MICCAI 2024. LNCS*, vol. 15005, pp. 564 – 574. Springer, Cham (2024).
18. Pölsterl, S., Wolf, T.N., Wachinger, C.: Combining 3D image and tabular data via the dynamic affine feature map transform. In: de Bruijne, M., et al. (eds.) *MICCAI 2021. LNCS*, vol. 12905, pp. 688 – 698. Springer, Cham (2021).
19. Bie, Y., Luo, L., Chen, Z., Chen, H.: XCoOp: explainable prompt learning for computer-aided diagnosis via concept-guided context optimization. In: Linguraru, M.G., et al. (eds.) *MICCAI 2024. LNCS*, vol. 15012, pp. 773 – 783. Springer, Cham (2024).
20. Diao, Y., Fang, H., Yu, H., et al.: Prior-guided prototype aggregation learning for Alzheimer’s disease diagnosis. In: Gee, J.C., et al. (eds.) *MICCAI 2025. LNCS*, vol. 15974, pp. 483–492. Springer, Cham (2025).
21. Kovac, E., Carlsson, S.V., Lilja, H., Hugosson, J., Kattan, M.W., Holmberg, E., Stephenson, A.J.: Association of baseline prostate-specific antigen level with long-term diagnosis of clinically significant prostate cancer among patients aged 55 to 60 years: A secondary analysis of a cohort in the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial. *JAMA Netw Open* 3(1), e1919284 (2020).
22. Hara, K., Kataoka, H., Satoh, Y.: Can spatiotemporal 3D CNNs retrace the history of 2D CNNs and ImageNet? In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2018)*, pp. 6546–6555 (2018).
23. Cao, W., Mirjalili, V., Raschka, S.: Rank consistent ordinal regression for neural networks with application to age estimation. *Pattern Recognit. Lett.* 140, 325–331 (2020).
24. Saha, A., Bosma, J.S., Twilt, J.J., et al.: Artificial intelligence and radiologists in prostate cancer detection on MRI (PI-CAI): An international, paired, non-inferiority, confirmatory study. *Lancet Oncol.* 25(7), 879–887 (2024).

25. Yu, Q., Zhou, W., Liu, H., et al.: Parameterized diffusion optimization enabled autoregressive ordinal regression for diabetic retinopathy grading. In: MICCAI 2025. LNCS, vol. 15974, pp. 450–460. Springer, Cham (2025).
26. Khosla, P., Teterwak, P., Wang, C., et al.: Supervised contrastive learning. In: Advances in Neural Information Processing Systems (NeurIPS), vol. 33, pp. 18661–18673 (2020).
27. Hao, X., Xu, H., Zhang, Q., et al.: Dual selective Gleason pattern-aware multiple instance learning for grade group prediction in histopathology images. In: MICCAI 2025. LNCS, vol. 15974, pp. 187–196. Springer, Cham (2025).
28. Duenias, D., Nichyporuk, B., Arbel, T., Riklin Raviv, T.: Hyperfusion: A hypernetwork approach to multimodal integration of tabular and medical imaging data for predictive modeling. *Med. Image Anal.* 102, 103503 (2025).