# LKAN: LLM-BASED KNOWLEDGE-AWARE ATTENTION NETWORK FOR CLINICAL STAGING OF LIVER CANCER: SUPPLEMENTAL MATERIAL

Our experiments are all based on Chinese radiology reports. The English translation in the main text is only for illustration. The corresponding explanations (figures or tables) from English to Chinese are as follows:

**Dataset Description**     The original dataset includes two parts: radiology reports and their corresponding staging (as shown in Table 1). The radiology report contains descriptions of lesions related to CSoLC, from which clinicians extract relevant information and compare it with the rules of CNLC one by one to stage each patient and develop subsequent treatment plans (hallmarks are detailed in the last column of Table 1). Our proposed method aims to learn from the experience of clinical doctors in first class hospitals and make staging judgments on radiology reports.

**Table 1.** Example sentences (Chinese version).

| Sentences | Staging | Hallmark(s) |
|---|---|---|
| 肝脏边缘欠光整，肝叶比例失调，肝裂增宽。肝S4及S5/6见两个类圆形低密度影，大小分别约31mm×28mm×28mm、33mm×33mm×32mm，增强动脉期不均匀强化，强化程度接近或稍低于邻近肝实质，门脉期强化程度减退，邻近肝静脉稍受压。 | IIa | Size of tumors $> 3$ (cm) Number of tumors $> 1$ |
| 平扫肝见多发大小不等、边界不清略低密度灶，双期增强扫描后，动脉期病灶周边部见多个小结节状强化灶，密度高于肝接近同层主动脉，门脉期病灶内强化灶范围扩大并逐渐向中央充填，最大者位于S5，大小约50mm×48mm。余肝实质未见异常强化灶。肝内、外胆管无扩张，肝门结构清晰。 | IIb | Size of tumors $> 3$ (cm) |
| S6见类圆形稍低密度影，直径约18mm，增强扫描动脉期呈等密度，门脉期呈低密度。门脉左支见卵圆形充盈缺损影，增强扫描轻度强化。 | IIIa | Vascular Invasion |
| 门静脉主干及其分支未见充盈缺损。肝内外胆道无扩张。肝门区可见一个增大的淋巴结，大小约为26.6mm×21.4mm，增强扫描后较明显强化。 | IIIb | Extrahepatic Metastasis |

The data format provided by the clinicians is as follows: the reports and their corresponding staging are shown in the first two columns of the table. We list the hallmarks in the last column to explain how clinicians summarize information related to CSoLC from radiology reports based on CNLC.

From Fig.2, it can be seen that the radiology report is lengthy, which increases the burden on doctors. We preprocess the report content to remove information unrelated to liver cancer, and use the content of "imaging findings" and "imaging diagnosis" to construct domain-specific word vectors (as shown in Fig.2(a)). Cosine similarity is used to calculate the adjacency relationship between different word vectors, and we can see that words related to stage-IIIa are close to each other (as shown in Fig.2(b)).

**Visualization**
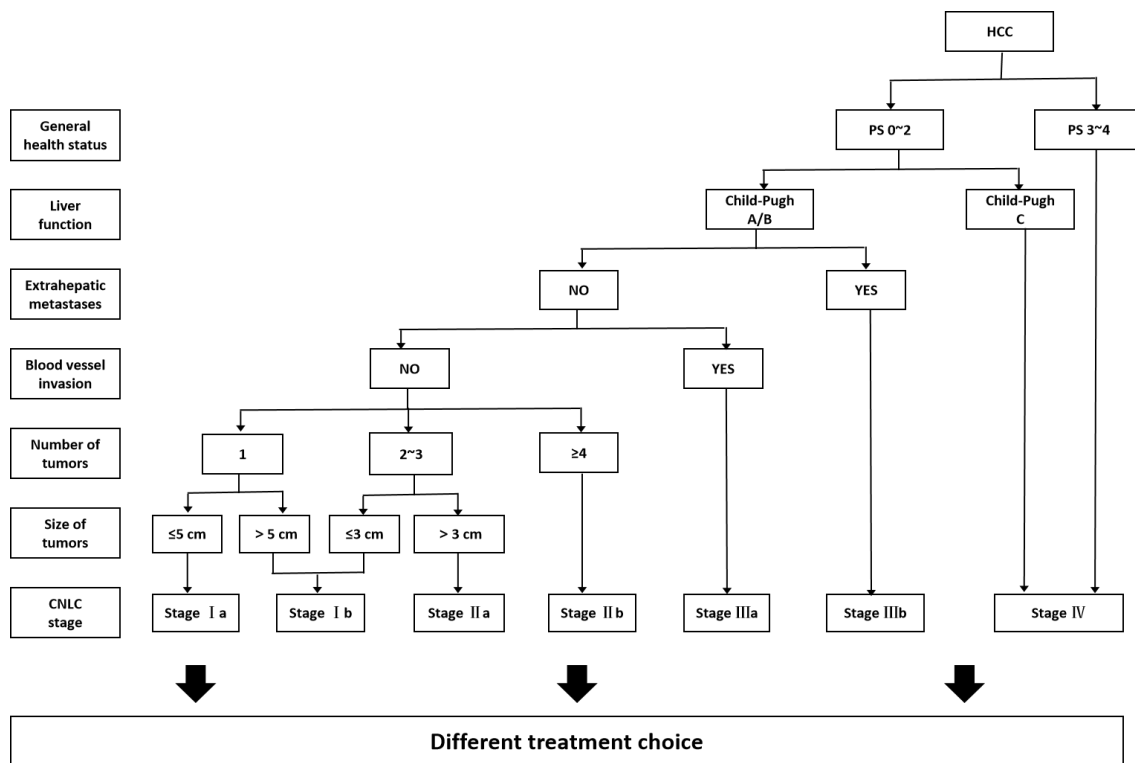
**Fig. 1.** China clinical staging and treatment pathway for PLC.
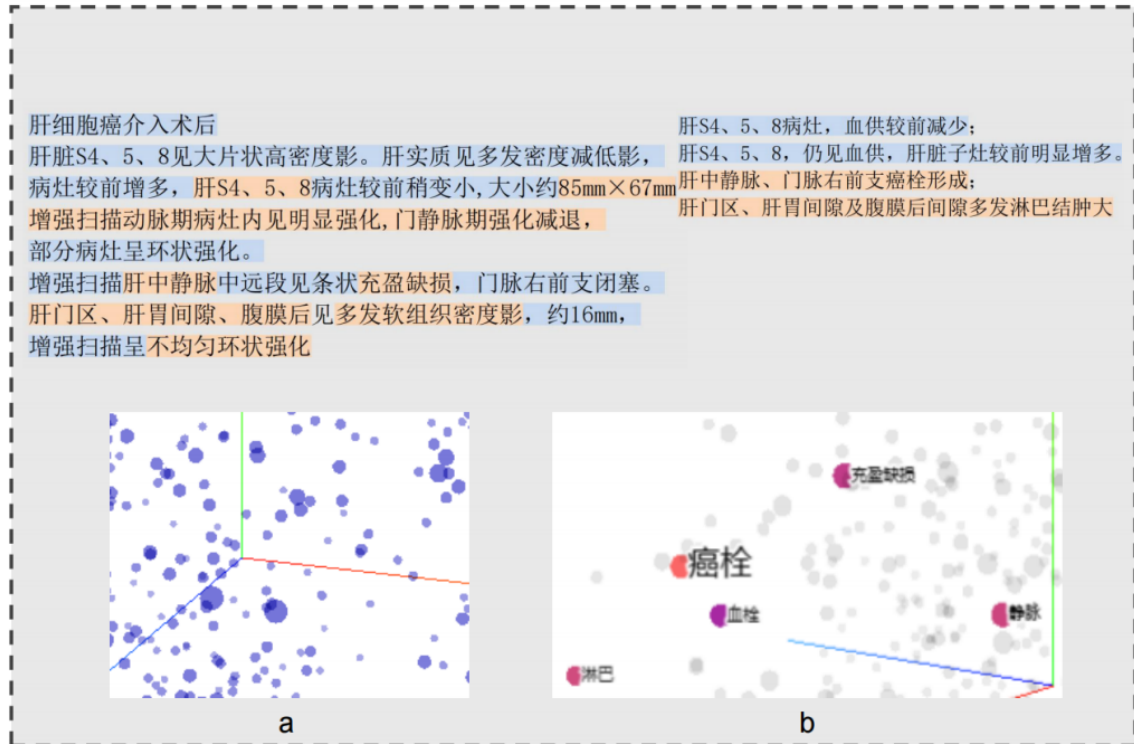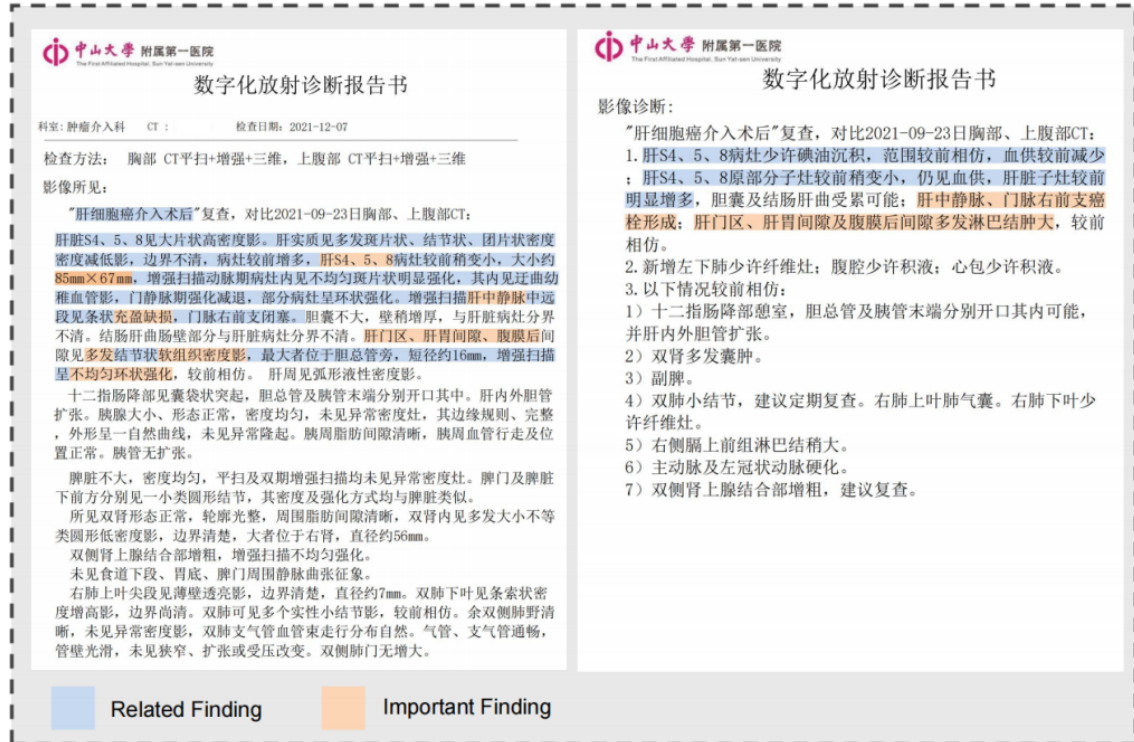
**Fig. 2.** An example of Chinese radiology reports (Chinese version). Blue blocks represent normal findings; Orange blocks represent relevant findings; Others indicate that they are not related to CSoLC. The distribution of pre-training word vectors is shown in (a). (b) showing spatial relationships of similar words.
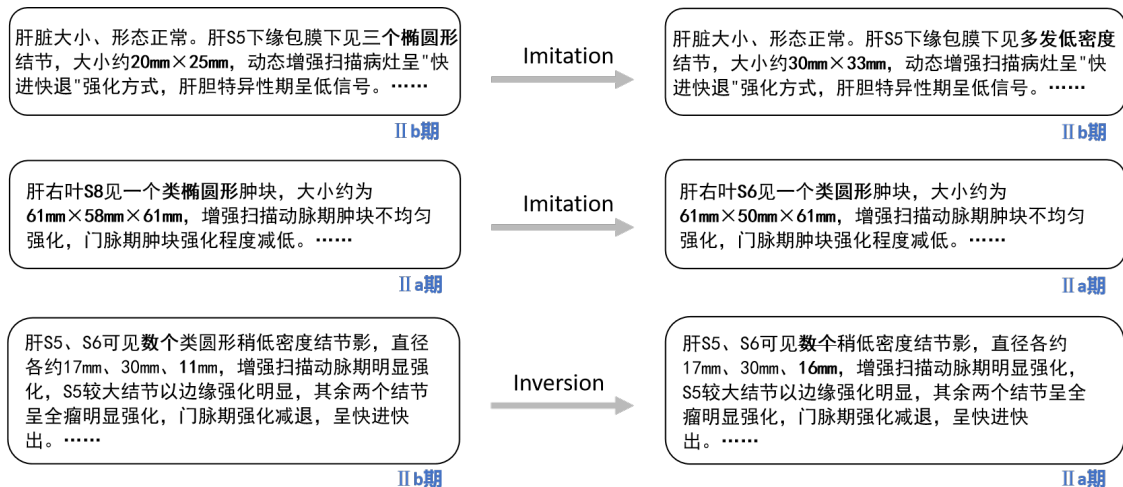
肝脏大小、形态正常。肝S5下缘包膜下见三**个椭圆形**结节，大小约20mm×25mm，动态增强扫描病灶呈"快进快退"强化方式，肝胆特异性期呈低信号。……

Ⅱb期

→ Imitation →

肝脏大小、形态正常。肝S5下缘包膜下见**多发低密度**结节，大小约30mm×33mm，动态增强扫描病灶呈"快进快退"强化方式，肝胆特异性期呈低信号。……

Ⅱb期

肝右叶S8见一个**类椭圆形**肿块，大小约为61mm×58mm×61mm，增强扫描动脉期肿块不均匀强化，门脉期肿块强化程度减低。……

Ⅱa期

→ Imitation →

肝右叶S6见一个**类圆形**肿块，大小约为61mm×50mm×61mm，增强扫描动脉期肿块不均匀强化，门脉期肿块强化程度减低。……

Ⅱa期

肝S5、S6可见**数个**类圆形稍低密度结节影，直径各约17mm、30mm、11mm，增强扫描动脉期明显强化，S5较大结节以边缘强化明显，其余两个结节呈全瘤明显强化，门脉期强化减退，呈快进快出。……

Ⅱb期

→ Inversion →

肝S5、S6可见**数个**稍低密度结节影，直径各约17mm、30mm、16mm，增强扫描动脉期明显强化，S5较大结节以边缘强化明显，其余两个结节呈全瘤明显强化，门脉期强化减退，呈快进快出。……

Ⅱa期

**Fig. 3.** Examples of "inversion" and "imitation" strategies (Chinese version).



**Fig. 4.** Visualization of attention weights (Chinese version).



**Fig. 5.** Visualization of the influence of redundant information on the classification results. (Chinese version).

4