



武汉大学

WUHAN UNIVERSITY

基于大模型的跨模态古诗创作

答辩人：张志东 软件工程 2021032111480

指导老师：朱卫平

武汉大学 计算机学院 本科毕业答辩

WUHAN UNIVERSITY
School of Computer Science



01 选题背景与意义

Background and Significance of Topic

03 实验结果与分析

Experiment Results and Analysis

基于大模型的跨模态古诗创作

研究思路与系统架构 02

Research Thought and System Architecture

论文总结与展望 04

Summary and Prospect of the Thesis



1.1 选题的背景与意义

古诗生成

独特的韵律规则与
艺术价值
中华优秀传统文化
再创造

大模型技术

强大的语言理解与
生成能力
更高质量的机器古
诗创作



1.2 现有挑战



现有方案局限于文本或图像的单一模态输入，难以捕捉用户完整的创作意图

- 文本输入：基于*RNN*进行韵律约束，生成的古诗停留在形式语义，无法触及典故、意象等更深层的文化内涵，缺乏文化深度与艺术性。
- 图像输入：基于*CNN*的接收图像输入，生成的古诗在情感主题等方面存在波动，不具备文本输入能够具体描述要求的优势



1.2 现有挑战



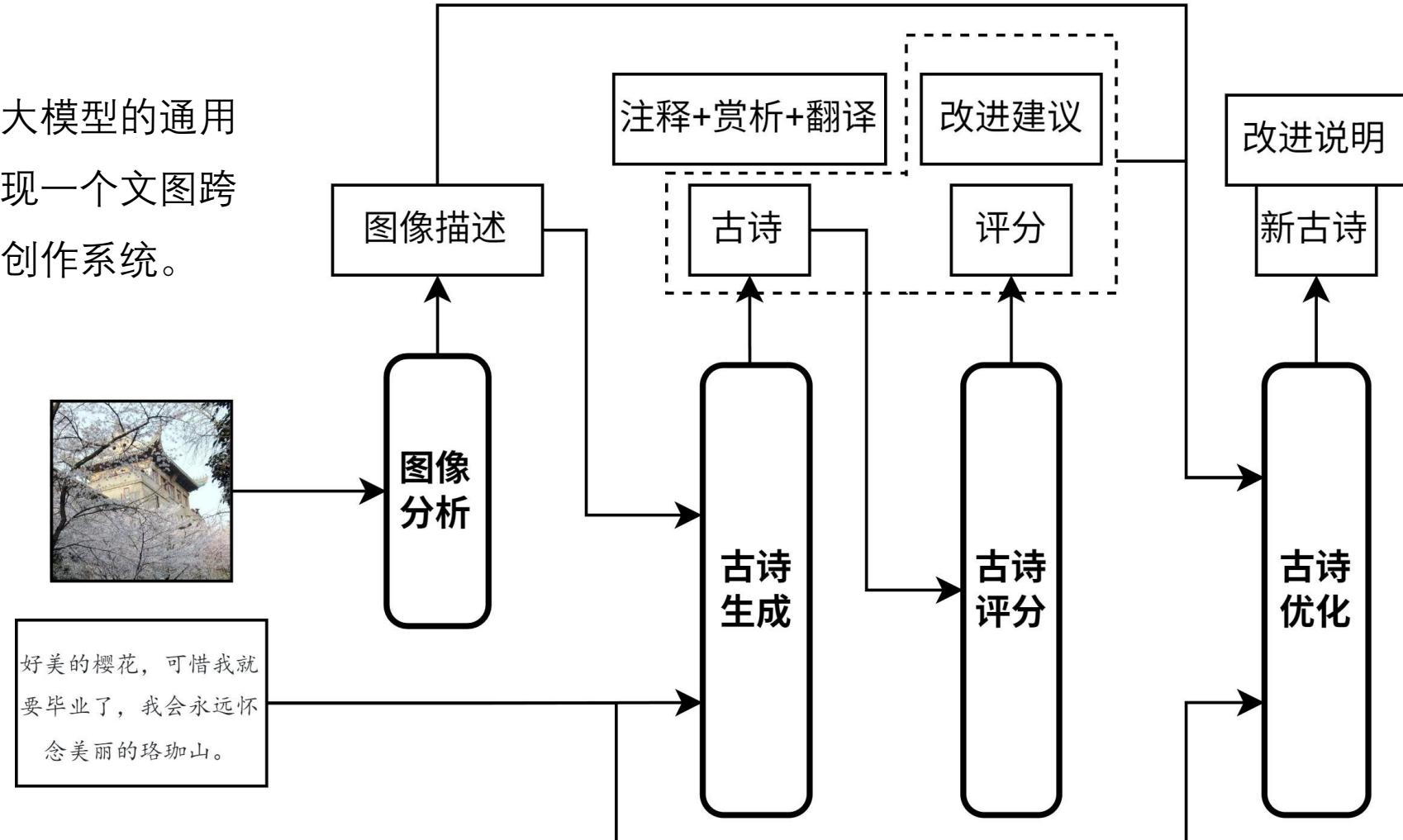
现有方案大多缺乏对生成过程和结果的可解释机制，对古诗的评价也缺乏可解释性。

- **过程上**：用户难以理解系统是**如何**从输入信息中提取出关键信息并**生成**古诗的。
- **结果上**：现有的系统也往往只提供了生成古诗的文本，而没有进一步**解释结果**，如韵律、意象等
- **反馈上**：现有的系统也缺乏对古诗**质量评价方法**的详细说明，用户需要具备较高的文学素养以理解和甄别对古诗质量的评价。



2.1 研究思路

本选题旨在利用大模型的通用能力，设计并实现一个文图跨模态的中文古诗创作系统。

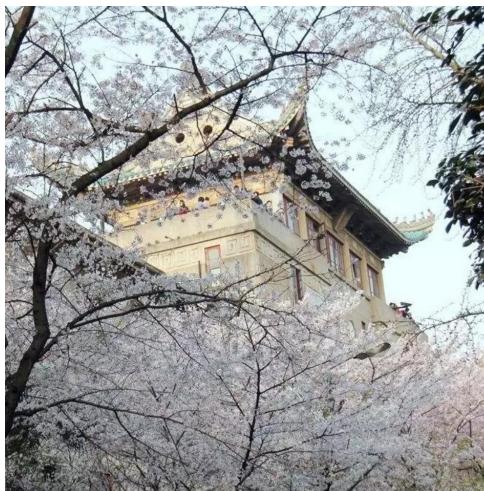




2.2 系统架构 – 图像分析

为处理输入的图像模态，需要使用跨模态能力优良的视觉与文本编码模型，使其能够：

1. 为图像生成描述文本，对图像中的物体特征、空间信息和整体氛围进行充分描述。
2. 识别出图片中的文化符号和隐含的情感信息，与中国传统文化要素产生勾连，便于后续古诗生成



DeepSeek-VL2: 这张图片展示了一座传统中式建筑，其屋顶装饰着精美的绿色琉璃瓦，显得庄重典雅。建筑主体为浅黄色，四周有栏杆围绕，似乎有游客站在阳台上欣赏风景。前景是繁茂的樱花树，粉白色的花朵覆盖了大部分画面，营造出浪漫而宁静的氛围。天空清澈，与花朵的柔和色彩形成鲜明对比，增添了画面的层次感。整幅画面让人感受到春天生机勃勃的气息，同时也透露出一种悠闲和诗意的情感。

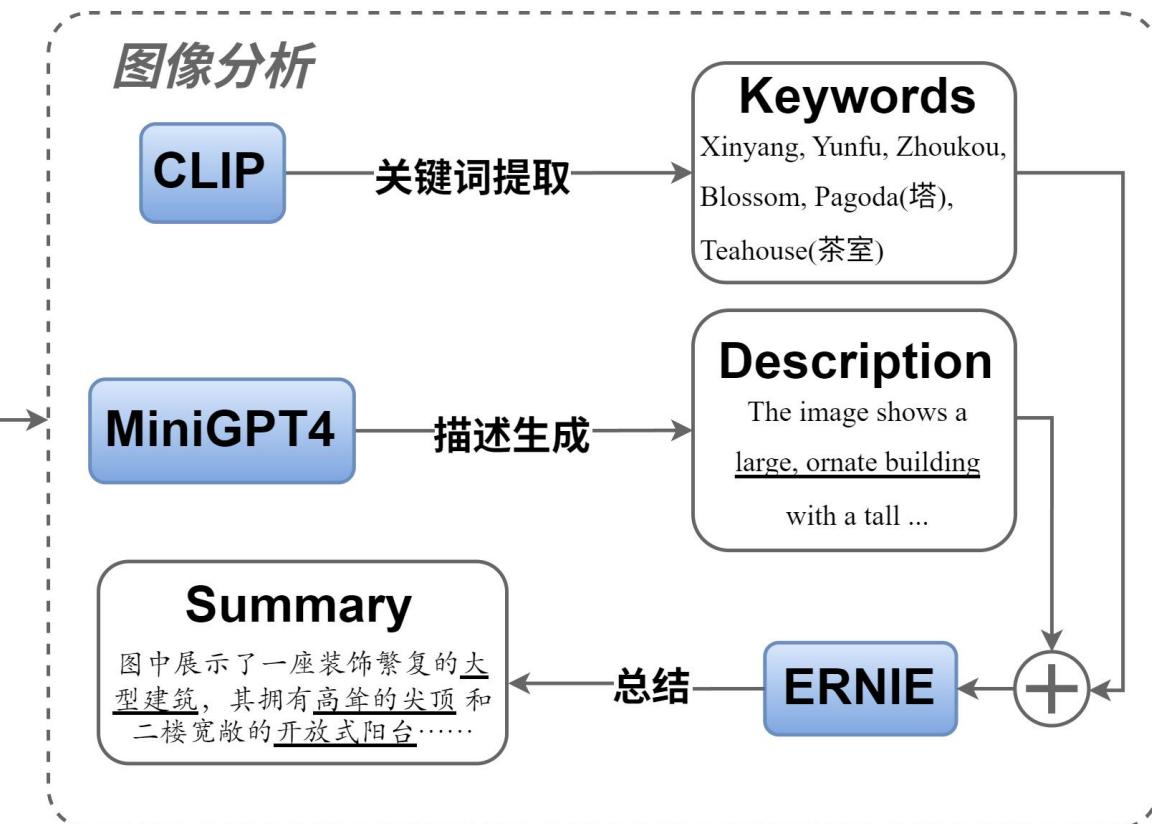


2.2 系统架构 – 图像分析

先前的方案存在CLIP 误读文化符号和MiniGPT4直译式描述的问题，最后的结果停留在缺失文化内涵的直译文本。



因此直接使用中文语境下的图像理解模型。

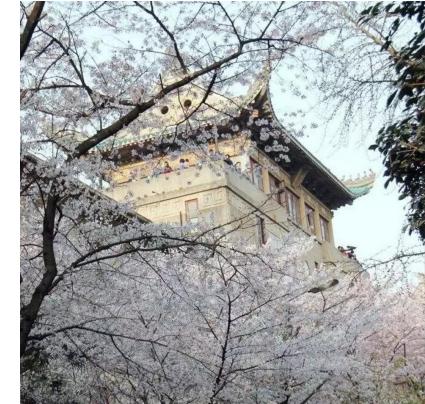




2.2 系统架构 – 图像分析

对比DeepSeek-VL2和Qwen2.5-VL-3B两个模型的表现，发现前者在色彩意象与空间层次上表现更好，因此选择DeepSeek-VL2用于图像分析。

- 色彩：前者的输出包含有“绿色琉璃瓦”、“浅黄建筑”、“粉白 樱花”等精确的色彩描述，而后者只模糊地提及“黄色和绿色”。
- 空间层次：前者通过“前景樱花-中景建筑-远景天空”的立体构图清晰地描述出了图像中的立体信息，而后者却丢失了这份空间关系的描述。



DeepSeek-VL2: 这张图片展示了一座传统中式建筑，其屋顶装饰着精美的绿色琉璃瓦，显得庄重典雅。建筑主体为浅黄色，四周有栏杆围绕，似乎有游客站在阳台上欣赏风景。前景是繁茂的樱花树，粉白色的花朵覆盖了大部分画面，营造出浪漫而宁静的氛围。天空清澈，与花朵的柔和色彩形成鲜明对比，增添了画面的层次感。整幅画面让人感受到春天生机勃勃的气息，同时也透露出一种悠闲和诗意的情感。



Qwen2.5-VL-3B: 这张图片展示了一座传统的中国建筑，可能是寺庙或宫殿。建筑的屋顶是典型的翘角设计，颜色为黄色和绿色。建筑周围有许多盛开的樱花树，花朵呈现出粉红色，与建筑的色彩形成了美丽的对比。树木的枝叶繁茂，覆盖了大部分画面。在建筑的阳台上，可以看到一些人在欣赏风景。整体氛围宁静而美丽，给人一种春天的感觉。

2.3 系统架构 – 古诗生成

为了优化用户体验、增强系统的可解释性，需要生成意象丰富的高质量古诗，并给出注释、赏析、白话文翻译等解释性文本。

珞珈山下琼英绽，朱阁檐前翠影摇。
 万点飞花迷画槛，千重香雪覆虹桥。
琉璃映日浮金阙，玉砌连云接碧霄。
 曲径曾携书卷过，长亭忍看落英飘。
 东君不解离人恨，更遣残红逐晚潮。
 粉泪暗随春水逝，青衿空对暮山遥。
 他年若问芳菲事，犹记晴川鹤影娇。

1. 琼英：取自《诗经·齐风》“尚之以琼英乎而”，既指樱花如玉，又喻学子才俊
2. 朱阁：对应图片中浅黄色建筑，化用苏轼“转朱阁”意象，暗喻学术殿堂
3. 虹桥：既实指校园拱桥，又暗用《东京梦华录》虹桥典故，喻求学之路
4. 琉璃：特指图片中绿色琉璃瓦，呼应白居易“琉璃波面月”的华美意象
5. 玉砌：化用李煜“雕栏玉砌应犹在”，将现代建筑古典诗化
6. 东君：司春之神，此处双关东风与毕业季的时令特征
7. 青衿：源自《诗经》“青青子衿”，既指学子装束，又喻求知初心
8. 晴川鹤影：化用崔颢“晴川历历汉阳树”，“鹤”取《诗经》“鹤鸣于九皋”意象，喻志向高远
9. 香雪：借李商隐“匝路亭亭艳”咏樱花之意，特指白色樱花如雪
10. 画槛：对应图片中栏杆意象，暗用“画栋朝飞南浦云”的滕王阁典故

珞珈山下樱花绽如琼瑶，朱红楼阁前翠影正轻摇。
 万点飞花迷蒙了雕栏画柱，千层香雪压弯了求学虹桥。
琉璃瓦映日似浮动金殿，白石阶入云直通向碧霄。
 曾捧书卷走过这林荫曲径，长亭外怎忍看落英飘摇。
 春神不懂人间别离苦楚，偏让残红追着晚潮远漂。
 粉泪默默随春水流逝去，青衫空对暮色里的山遥。
 若问当年花开烂漫旧事，定记得晴川上鹤影正娇。

这首七言排律以武大樱花为情感载体，通过工丽的对仗与典雅的意象，构建出毕业离别的双重意境。

首联以“琼英”“翠影”点题，既暗合《诗经》“尚之以琼英乎而”的典故，又呼应图片中琉璃瓦与樱花的色彩映照。颔联“万点飞花”与“千重香雪”形成空间纵深感，化用岑参“千树万树梨花开”的意境而翻出新意。颈联转写建筑，“琉璃映日”直取图片意象，“玉砌连云”暗用李煜“雕栏玉砌”之典，将现代学府与传统意象完美融合。尾联“晴川鹤影”既化用崔颢“晴川历历汉阳树”，又暗喻学子如鹤高飞，留下余韵悠长的想象空间。

全诗严守平水韵，中二联“迷画槛”“对”“覆虹桥”，“浮金阙”“对”“接碧霄”，工稳中见灵动，以物象变迁写时光流逝，最终在“鹤影”的意象中完成对母校记忆的诗意图定格。



2.3 系统架构 – 古诗生成

测试多种大语言模型对典故意象的运用能力，发现R1模型显著更优。

ERNIE-4.0:

珞珈山：指用户所怀念的校园所在地，此处借指校园。
樱如雪：形容樱花盛开时如雪般洁白美丽。
绿瓦黄墙：描述传统中式建筑的屋顶绿色琉璃瓦和浅黄色墙体。
春华：春天的花，此处泛指春天的美景。
花满袖：形容花香扑鼻，似乎连衣袖都沾满了花香。……

DeepSeek-V3:

珞珈山：位于中国湖北省武汉市，武汉大学所在地，以樱花闻名。
樱雪：形容樱花盛开如雪。
翠微：山色青翠。
琉璃瓦：传统中式建筑屋顶的装饰材料，绿色琉璃瓦显得庄重典雅。
黄墙：传统中式建筑的主体颜色，浅黄色。……

DeepSeek-R1:

琼英：取自《诗经·齐风》“尚之以琼英乎而”，既指樱花如玉，又喻学子才俊
朱阁：对应图片中浅黄色建筑，化用苏轼“转朱阁”意象，暗喻学术殿堂
虹桥：既实指校园拱桥，又暗用《东京梦华录》虹桥典故，喻求学之路
琉璃：特指图片中绿色琉璃瓦，呼应白居易“琉璃波面月”的华美意象
玉砌：化用李煜“雕栏玉砌应犹在”，将现代建筑古典诗化……

更多是对景色的直观描述，缺乏对文化典故的深度挖掘和巧妙运用，尤其体现在输出的“注释”部分

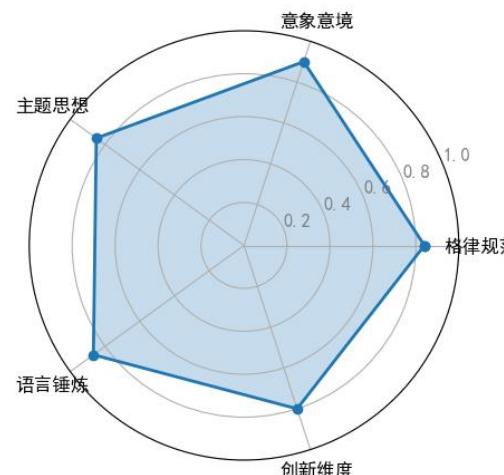




2.4 系统架构 – 古诗评价

为合理评价古诗质量，设计评分体系并调用DeepSeek-R1模型打分，给出评分和针对性的优化建议。

珞珈山下琼英绽，朱阁檐前翠影摇。
万点飞花迷画槛，千重香雪覆虹桥。
琉璃映日浮金阙，玉砌连云接碧霄。
曲径曾携书卷过，长亭忍看落英飘。
东君不解离人恨，更遣残红逐晚潮。
粉泪暗随春水逝，青衿空对暮山遥。
他年若问芳菲事，犹记晴川鹤影娇。



[评分]

1. 格律规范 (21/25)

1.1 平仄音韵 (8/10)：“琼英绽”三平尾（琼-平、英-平、绽-仄，按平水韵“绽”属上声为仄，实为三平调；“忍看落英飘”存在孤平现象）

1.2 对仗工稳 (8/10)：“琉璃映日”对“玉砌连云”器物建筑对仗精妙，但“曲径曾携”与“长亭忍看”动词结构稍欠工整

1.3 押韵协调 (5/5)：全诗押平水韵二萧部（摇、桥、霄、飘、潮、遥、娇）如编钟贯珠

2. 意象意境 (27/30)

2.1 古典运用 (18/20)：“晴川鹤影”化用崔颢典而不露，将“珞珈山”地域特征融入古典语境如盐入水

2.2 意境层次 (9/10)：从实景（琼英绽）到追忆（书卷过）再到时空穿越（鹤影娇），构建三重意境如敦煌飞天飘带

3. 主题思想 (17/20)

3.1 情感真挚 (11/12)：“青衿空对暮山遥”将求学记忆与离别惆怅交织，如吴带当风

3.2 思想传承 (6/8)：“东君不解”暗合《楚辞》司春之神原型，但未突破传统伤春范式

4. 语言锤炼 (13/15)

4.1 凝练度 (7/8)：“万点飞花迷画槛”数字量词精准如界画，唯“忍看”稍显直露

4.2 典雅度 (6/7)：“粉泪”承李煜“胭脂泪”意象，“碧霄”接刘禹锡“晴空一鹤”语境

5. 创新性 (8/10)

5.1 守正出新 (8/10)：“珞珈山”地理符号与“金阙”仙家意象融合，如唐三彩吸收胡风

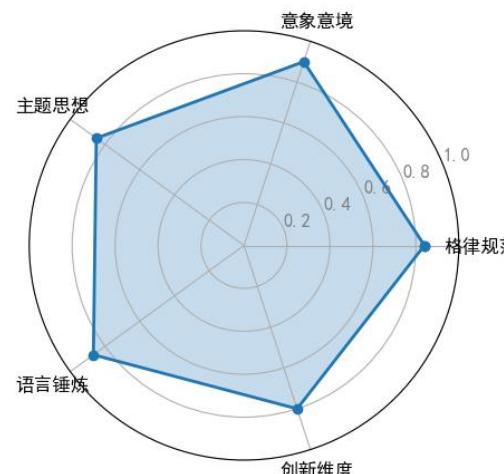
总分: 86/100



2.4 系统架构 – 古诗评价

为合理评价古诗质量，设计评分体系并调用DeepSeek-R1模型打分，给出评分和针对性的优化建议。

珞珈山下琼英绽，朱阁檐前翠影摇。
万点飞花迷画槛，千重香雪覆虹桥。
琉璃映日浮金阙，玉砌连云接碧霄。
曲径曾携书卷过，长亭忍看落英飘。
东君不解离人恨，更遣残红逐晚潮。
粉泪暗随春水逝，青衿空对暮山遥。
他年若问芳菲事，犹记晴川鹤影娇。



[优化建议]

1. 格律规范

1.1 平仄音韵:

- 首联“琼英绽”改“琼瑶碎”（平平仄），既避三平调又暗含陆机《文赋》“玉卮无当，虽宝非用；琼瑶满案，终难为佩”的喻象

- “忍看落英飘”调整为“忍顾乱红凋”（仄仄仄平平），参考秦观“忍顾鹊桥归路”句式

2. 意象意境

2.1 古典运用:

- “晴川鹤影”可深化为“晴川历历汉阳树”的今昔对照，加入“残碑”意象与“鹤影”形成文明传承的张力

- “琉璃映日”建议改为“琉璃承露”，接续李贺“琉璃钟，琥珀浓”的瑰丽想象，增强晨昏时序感

3. 主题思想

3.2 思想传承:

- 尾联注入《论语》“浴乎沂，风乎舞雩”的哲思，将“芳菲事”升华为“曾点之志”，改“犹记晴川鹤影娇”为“独对春沂咏雩娇”，使离愁转为文化坚守

4. 语言锤炼

4.1 凝练度:

- “朱阁檐前”精炼为“朱阑十二”，化用李白“解释春风无限恨，沉香亭北倚阑干”的典故密度

- “香雪覆虹桥”调整为“香阵没星轺”，借“星轺”（使者车驾）暗示人生旅途，增强叙事性

5. 创新性

在“残红逐晚潮”处植入现代意象：“无人机影掠江皋”，形成古典送别（兰舟）与当代离别（无人机）的蒙太奇，呼应李商隐“却话巴山夜雨时”的时空折叠技法



2.4 系统架构 – 古诗评价

对先前设计的评分体系进行重构，解决了三个核心问题：

- **评分标准细化**：对每个子维度设置明确的**分数段**，提高给分区分度；
- **维度解释充分**：为每个子维度的每个分数段设置典型示例说明；
- 重构**维度逻辑**：重构之前互有重叠的维度，提高体系合理性；

1. 结构与形式 (Structural and Formal Aspects) - 总分 10 分

诗歌类型：最高 5 分。评估是否符合指定类型（如律诗、绝句）的基本结构和规则。

韵律规则：最高 5 分。分析诗歌的韵律是否规整，是否符合传统韵律规则。



1. 格律规范 Prosodic Norms (25 points)

1.1 平仄音韵 Tonal Prosody (10 points)

9-10: 完全符合唐体格律（例：杜甫《登高》“风急天高猿啸哀，渚清沙白鸟飞回”平仄严谨）

7-8: 个别拗句但有救（例：王维《终南别业》“行到水穷处”第三字拗，第四字救）

5-6: 三平尾/三仄尾不超过两处（例：韦应物《滁州西涧》“独怜幽草涧边生”三平尾）

0-4: 严重失律（例：打油诗体）

1.2 对仗工稳 Antithetical Structure (10 points)

9-10: 工对+借对精妙（例：李商隐《锦瑟》“庄生晓梦迷蝴蝶，望帝春心托杜鹃”）.....



2.4 系统架构 – 古诗评价

而为了避免较复杂的评分体系干扰提示词的其他描述，参考**Few-shot**框架给定两个输入输出的范例，确保模型按要求输出。

下面是对创作古诗的详细评分规则，我将提供古诗文本给你，请你严格基于以下进行评价和打分。在完成评分后，我还需要你提供对古诗文本的具体改进建议，同样依照这些角度来一一说明，特别是得分较低的角度。注意输出只包含“评分”与“改进建议”两部分，且须严格遵循格式要求。

<评分体系>

【输入】 【输出】

【输入】 【输出】



2.4 系统架构 – 古诗评价

而为了避免较复杂的评分体系干扰提示词的其他描述，参考Few-shot框架给定两个输入输出的范例，确保模型按要求输出。

下面是对创作古诗的详细评分规则，我将提供古诗文本给你，请你严格基于以下进行评价和打分。在完成评分后，我还需要你提供对古诗文本的具体改进建议，同样依照这些角度来一一说明，特别是得分较低的角度。注意输出只包含“评分”与“改进建议”两部分，且须严格遵循格式要求。

<评分体系>

【输入】 【输出】

【输入】 【输出】

此外，利用自动度量指标作为古诗质量的辅助参考。

- **BLEU、ROUGE**: 计算参考文本与候选文本的n-gram重叠程度
- **Distinct**: 计算古诗中独特的n-gram占比
- **Similarity**: 基于词向量计算古诗内前后句子的语义相似度，如同一联内前后句的相似度**Sim_intra**、不同联之间的相似度**Sim_inter**

Sim_intra

珞珈山下琼英绽，朱阁檐前翠影摇。
万点飞花迷画槛，千重杳雪覆虹桥。
琉璃映日浮金阙，玉砌连云接碧霄。
曲径曾携书卷过，长亭忍看落英飘。
东君不解离人恨，更遣残红逐晚潮。
粉泪暗随春水逝，青衿空对暮山遥。
他年若问芳菲事，犹记晴川鹤影娇。

Sim_inter



2.5 系统架构 – 古诗优化

为有效改进生成的古诗，需要参考改进意见与参考原古诗的评分，同时应考虑文本图像输入，以避免偏离用户意图。产生优化后古诗与改进说明。

珞珈山下琼英绽，朱阁檐前翠影摇。
万点飞花迷画槛，千重香雪覆虹桥。
琉璃映日浮金阙，玉砌连云接碧霄。
曲径曾携书卷过，长亭忍看落英飘。
东君不解离人恨，更遣残红逐晚潮。
粉泪暗随春水逝，水逝生认空对首小泣。
他年若问芳



珞珈山下琼瑶碎，朱阑十二翠云飘。
万点飞花迷玉砌，千重香阵没星轺。
琉璃承露浮金阙，玉砌连云接碧霄。
曲径曾携黄卷过，长亭忍顾乱红凋。
东君未解青衿恨，犹遣残霞逐晚潮。
粉泪暗随春水逝，素襟空对暮山遥。
残碑鹤影参商渡，机影江声日夜迢。
欲问蓬瀛何处是，春沂独咏舞雩娇。

86 → 90



2.5 系统架构 – 古诗优化

为有效改进生成的古诗，需要参考改进意见与参考原古诗的评分，同时应考虑文本图像输入，以避免偏离用户意图。产生优化后古诗与改进说明。

珞珈山下琼英绽，朱阁檐前翠影摇。
万点飞花迷画槛，千重香雪覆虹桥。
琉璃映日浮金阙，玉砌连云接碧霄。
曲径曾携书卷过，长亭忍看落英飘。
东君不解离人恨，更遣残红逐晚潮。
粉泪暗随春水逝，水逝素襟空对暮山遥。
他年若问芳



珞珈山下琼瑶碎，朱阑十二翠云飘。
万点飞花迷玉砌，千重香阵没星轺。
琉璃承露浮金阙，玉砌连云接碧霄。
曲径曾携黄卷过，长亭忍顾乱红凋。
东君未解青衿恨，犹遣残霞逐晚潮。
粉泪暗随春水逝，素襟空对暮山遥。
残碑鹤影参商渡，机影江声日夜迢。
欲问蓬瀛何处是，春沂独咏舞雩娇。

86 → 90

1. 平仄调整：“琼瑶碎”（平平仄）化解原三平尾，暗引陆机“琼瑶”喻指未琢之玉，喻学子未竟之志；“忍顾乱红凋”（仄仄仄平平）借易安“忍顾”词眼，使孤平转双拗救
2. 意象重构：“朱阑十二”化用李白沉香亭典故，以数字阑干喻求学四载；“香阵没星轺”借骆宾王“香阵冲天”与谢灵运“星轺阻辙”，喻青春壮志与宦途阻隔的冲突
3. 时空折叠：新增颈联“残碑鹤影参商渡，机影江声日夜迢”，以崔颢“晴川历历”对应珞珈残碑，无人机影接续李商隐“星桥铁锁”意象，构建古今送别对话场域
4. 哲思升华：尾联化《论语》“风乎舞雩”为“春沂独咏”，将樱花意象升华为“浴沂归咏”的文化坚守，使离愁转为对“曾点之志”的传承
5. 器物对仗：调整“琉璃承露”对“玉砌连云”，既保持建筑意象，又以“承露”接续李贺琉璃钟意象，暗喻学子承接文化甘露
6. 色彩体系：保留原诗“翠、金、碧、粉、青”色谱，新增“素襟”与“残霞”形成冷暖对照，暗合图片中琉璃瓦的冷绿与樱花的暖粉



2.5 系统架构 – 古诗优化

作为对比，若只提供原古诗和改进建议，会导致优化效果不佳，且输出内容偏离用户意图。

86 → 90

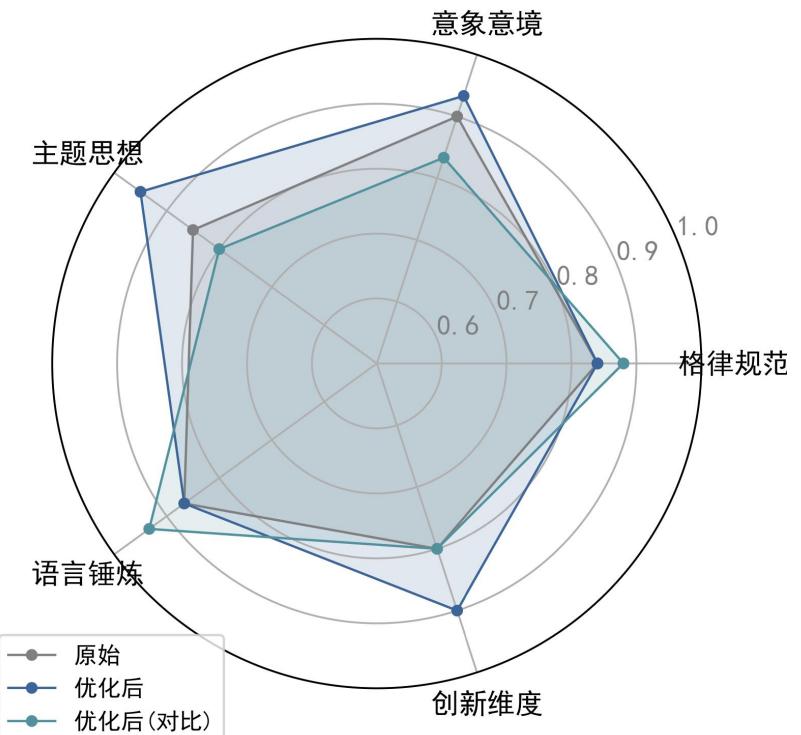
珞珈山下琼瑶碎，朱阑十二翠云飘。
万点飞花迷玉砌，千重香阵没星轺。
瑠璃承露浮金阙，玉砌连云接碧霄。
曲径曾携黄卷过，长亭忍顾乱红凋。
东君未解青衿恨，犹遣残霞逐晚潮。
粉泪暗随春水逝，素襟空对暮山遥。
残碑鹤影参商渡，机影江声日夜迢。
欲问蓬瀛何处是，春沂独咏舞雩娇。

优化后

86 → 85

珞珈山下琼瑶碎，朱阑十二翠云招。
万点飞花迷画舫，千重香阵没星轺。
瑠璃承露浮金阙，玉砌连霄接绛绡。
曲径曾携黄卷过，长亭忍顾乱红凋。
残碑犹刻汉阳树，鹤影空临楚泽苔。
东君未解离人恨，却遣春醒逐暮飙。
粉泪暗随青翰逝，玄裳独对紫烟缭。
无人机影掠江皋，兰棹声沉浦溆遥。
若问当年芳菲志，曾点春沂咏雩娇。

优化后（对比）





3.1 实验分析

为验证系统功能，设计了三项实验：

1. 基于白话文的古诗生成实验 —— 对比ERNIE-4.0与DeepSeek-R1模型生成古诗的表现
2. 评分功能实验 —— 检验系统评分功能、自动度量指标的有效性
3. 文图结合的古诗生产实验 —— 测试文图模态输入对生成结果的影响

3.2 实验分析 – 基于白话文的古诗生成实验

以《唐诗三百首》所有律诗绝句的白话文翻译为输入，基于**ROUGE**和**BLEU**计算模型输出与原古诗的相似性，为对比**DeepSeek-R1**与**ERNIE-4.0**两个模型生成古诗的表现。

测试发现，二者的输出与原文存在高度重合，计算出的分数也明显区别于现有工作的结果（如 $\text{BLEU-1}=0.168$, $\text{BLEU-2}=0.002$ ），考虑是原古诗包含于大模型训练语料，故无法得出结论。

表 4.2 白话文古诗生成实验结果 (**DeepSeek-R1**)

	BLEU		ROUGE		
	BL-1	BL-2	R-1	R-2	R-L
七言律诗	0.583599	0.432841	0.557353	0.349461	0.527267
七言绝句	0.559833	0.415325	0.562665	0.344502	0.540417
五言律诗	0.597726	0.432494	0.549401	0.337500	0.545573
五言绝句	0.523605	0.332445	0.495130	0.245845	0.471983
平均	0.575037	0.414674	0.546996	0.329415	0.529737

表 4.3 白话文古诗生成实验结果 (**ERNIE-4.0**)

	BLEU		ROUGE		
	BL-1	BL-2	R-1	R-2	R-L
七言律诗	0.765951	0.671162	0.770484	0.609392	0.743464
七言绝句	0.798136	0.702556	0.808346	0.649742	0.764375
五言律诗	0.691980	0.533006	0.670557	0.445450	0.646720
五言绝句	0.839312	0.764698	0.834125	0.726052	0.812500
平均	0.767420	0.658670	0.765072	0.596372	0.734993



3.3 实验分析 – 评分功能实验

为检验系统评分功能的有效性，收集第六届“诗词中国”传统诗词创作大赛的公开作品，利用专家评审出的奖项设置，与《唐诗三百首》中的作品、打油诗一同作为分层测试集。

表 4.7 纳入唐诗和打油诗的评分实验（系统评分）

	格律规范	意象意境	主题思想	语言锤炼	创新性	总分	样本
唐诗	0.932	0.934	0.931	0.930	0.853	0.924	98
一二三	0.768	0.818	0.839	0.818	0.821	0.810	28
优秀奖	0.762	0.805	0.824	0.798	0.797	0.796	91
打油诗	0.328	0.490	0.615	0.470	0.585	0.481	20

测试发现，系统评分的结果与真实结果完全一致（包括各维度与总分的比较结果），且不同组别的给分均值存在明显差异。



3.3 实验分析 – 评分功能实验

此外，也对自动度量指标进行了检验，发现BLEU与ROUGE两个指标表现较优，而体现用词多样性的Distinct则未体现独立反映质量的能力。

表 4.8 纳入唐诗和打油诗的评分实验（BLEU+ROUGE）

	BLEU		ROUGE		
	BL-1	BL-2	R-1	R-2	R-L
唐诗	0.905665	0.792163	0.125864	0.005379	0.157440
一二三	0.945848	0.781568	0.113968	0.004003	0.145957
优秀奖	0.934612	0.749390	0.107799	0.003346	0.143840
打油诗	0.893658	0.726545	0.095770	0.003127	0.137659

表 4.9 纳入唐诗和打油诗的评分实验（Similarity+Distinct）

	Similarity		Distinct	
	S-Intra	S-Inter	D-1	D-2
唐诗	0.678716	0.693137	0.854964	0.998430
一二三	0.661024	0.694880	0.885342	0.999165
优秀奖	0.670973	0.696803	0.879247	0.999832
打油诗	0.698042	0.738067	0.760565	0.972817

Similarity的结果与真实结果差异很大，其中**Sim_inter**的结果完全相反，考虑**指标设计问题**，如不同联之间的往往会切换主题，因而语义差异较大(**Sim_inter**)，或是选取的**词向量Bert-CCPoem**不适合此类下游任务。



3.4 实验分析 – 文图结合的古诗生成实验

为测试文图输入对生成结果的影响，使用 5 对图像和文本作为输入，在文、图、文图三种不同的输入组合下分别进行 10 次生成，计算系统评分均值。

表 4.10 文图古诗生成实验结果

	格律规范	意象意境	主题思想	语言锤炼	创新性	总分
文	0.862	0.901	0.883	0.871	0.850	0.878
图	0.862	0.899	0.878	0.869	0.832	0.875
图文	0.866	0.900	0.885	0.861	0.832	0.876

测试发现，文图模态输入对生成古诗质量的影响并不显著（差异均小于0.01）。古诗的质量更与模型能力与系统架构有关，而图文输入更多地用于提高用户体验。



3.4 实验分析 – 文图结合的古诗生成实验

日出时分，山间草甸被晨光染成金色，仿佛时间静止。这一刻的宁静，让人忘却尘世喧嚣，只想与这美景共存。



落花有意，流水无情，这满地的花瓣，像是春天的泪痕。它们曾是枝头的骄傲，如今却只能随风飘散，让人不禁感叹生命的短暂和无常。



.....

千嶂含烟衔夕照，金鳞万顷泛柔荑。
赩霞散绮浮仙袂，碧落流丹浣羽衣。
黛岫逶迤吞暮霭，青芜起伏接鸿熹。
云根欲堕瑶池色，风末犹传帝子矶。
野径斜曛迷鹿迹，天光倒影漾星玑。
凭高忽觉尘襟豁，醉倚青崖忘世机。

簇锦垂云倚壁开，飞琼碎玉漫空徊。
残垣著雨斑痕旧，冷砌堆香蝶魄哀。
晓露初凝鲛绡泪，斜阳半染蜀绡腮。
金铃已绝三更梦，罗袜空遗九曲苔。
飘转犹寻青帝诺，伶俜忍负紫霞杯。
鶗鴂声咽千枝血，蛱蝶魂销万点灰。
莫问芳心归哪处，东君自解送春来。



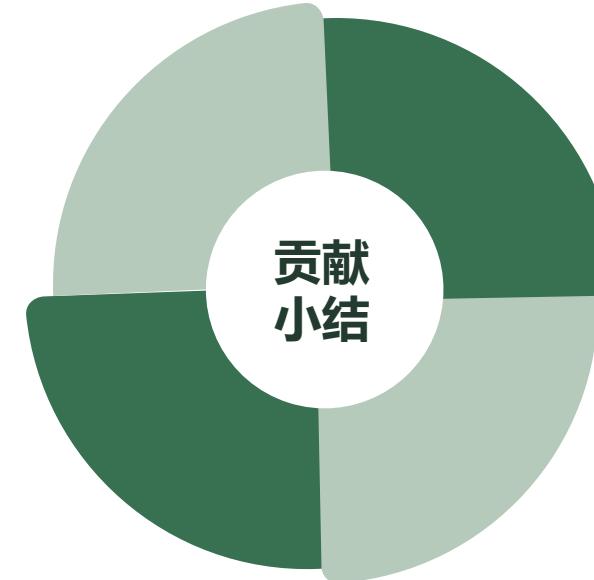
4.1 贡献总结

文图古诗生成

结合文图双模态来强化生成古诗与用户需求的契合度，输出既符合格律要求、又富有典故意象的高质量古诗

古诗评价

设计维度合理、标准明确的古诗评分体系，评分效果相比自动度量方法更稳定



结果可解释性

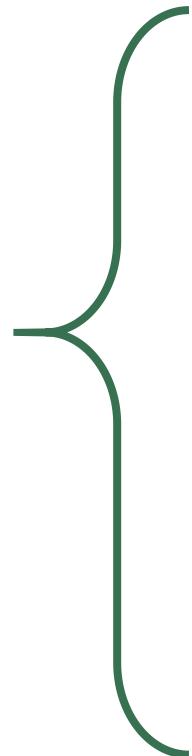
在生成古诗时，输出注释等用户友好的解释性文本，提高了系统输出结果的可解释性

古诗优化

结合用户输入和评价结果，实现了兼顾作品质量与用户意图的古诗优化。



4.2 不足

- 
- ① 本系统并未实现在向量层面上的文图融合，而是通过文本描述的方式进行模态间的转换，可能导致信息损失或语义偏差
 - ② 文图模态输入对生成过程的作用仍有待进一步验证，尤其在古诗形式质量上遇到瓶颈后。
 - ③ 评分机制的准确性十分依赖于大模型自身的质量，评分与优化的效果仍存在一定波动
 - ④ 系统推理时间较长



武汉大学
WUHAN UNIVERSITY

感谢倾听 请各位老师批评指正

基于大模型的跨模态古诗创作

答辩人：张志东 × 指导老师：朱卫平

WUHAN UNIVERSITY
School of Computer Science
28