

## 个人研究陈述 (Statement of Purpose)

宋有哲 (Youzhe Song)

申请项目：计算机科学博士 (PhD in Computer Science)

申请地区：北美 (North America)

申请学期：2026 年秋季 (Fall 2026)

### 1. 引言：我的研究兴趣与思考

当前人工智能模型在各项任务中取得了显著的成功，但它们的内部工作原理很大程度上仍是一个“黑盒”。我们知道模型能做什么 (what)，并在一定程度上、一些细节中知道它是如何做到 (how) 的，但缺乏对于整个认知过程、思考过程符合人类直觉的理解。我认为，通往更通用、更可靠 AI 的重要一步，在于理解并设计模型内部的信息处理机制。我的研究兴趣聚焦于两个核心问题：第一，我们能否在模型内部构建一个**符合人类直觉的、分步、有机的推理路径**？第二，我们如何让模型像人一样，**理解并对齐来自不同来源（如图像、文本）的信息**？

我的目标是探索如何在高维的隐空间 (latent space) 中，实现这两种能力的融合，让模型不仅能“看懂”和“读懂”世界，更能在此基础上进行清晰、可靠的推理。

### 2. 研究实践一：设计更鲁棒的特征表示 (CoReFace)

我的研究之旅始于对自监督对比学习的深刻向往。它不依赖昂贵的人工标注，而是从数据本身的结构中学习，这种优雅与潜力让我着迷。然而，现实的困境一个计算资源极为有限且每个学生研究方向都不一样的实验室环境迫使我放弃对算力蛮力的依赖，转而思考一个更根本的问题：如何通过精巧的**机制设计**来提升模型性能。

为此，我发现了人脸识别这个完美的“沙盒”。我洞察到它与对比学习在本质上的深刻相似性：两者都在高维空间中精心规划特征的分布几何。基于此，我提出了 CoReFace 框架，这是我对模型内部机制的一次“显式设计”。在人脸这种细粒度识别任务中，传统的图像增强会破坏关键身份信息。我当时的一个关键洞察是，Dropout 本质上是一种**特征层面的随机扰动**，它能在不破坏图像语义的前提下，为对比学习提供必要的视角。通过引入一个由对比学习引导的正则项，我主动地去“雕刻” (sculpting) 特征空间的几何结构。最终，该方法成功将正负样本对的**相似度边界** (similarity margin) **提升了 15%**。在这项工作中，我主导了从识别挑战、设计方案到独立验证与论文撰写的完整生命周期，完成了从“学生”到“**独立研究者**”的身份认知转变 [1]。

### 3. 研究实践二：构建异构数据的统一处理框架 (QGFace)

完成 CoReFace 后，我渴望研究能与真实世界产生更直接的联系。这个想法在我翻看手机相册时得到了印证：相册自带的 AI 人脸聚类功能，在处理我的家庭合照时表现糟糕。我意识到，现实世界中因构图、光线等原因产生的低质量数据，与我们常用的数据集有着重大的区别。

QGFace 的**单编码器** (single-encoder) **架构**，正是我为此设计的、内置了智能“**信息调度系统**”的解决方案。它模拟一种“分诊” (triage) 过程：高质量数据走分类通道，低质量数据走对比通道，并用梯度截断来防止信息污染。当发现对比学习因样本对数量不足而效果不彰时，我设计了一个**动态更新的编码池**，显著增强了其性能。为了追赶截稿日期，我全身心投入其中。那段时间，我的一天被切分成数个专注的循环，GPU 的工作时间才是我的休息时间。这段经历对我而言并非负担，而是对我研究热情的终极“**压力测试**”：它让我确信，驱动我的不是外部的期望，而是解决难题时发自内心的智识乐趣。最终，QGFace 在几乎不牺牲高质量数据性能（仅有 0.3% 的**微小性能权衡**）

的情况下，在低质量数据集上达到了 SOTA，证明了我的研究哲学能指导我设计出兼具优雅与鲁棒性的实用方案 [2]。

#### 4. 探索与聚焦：在实践中明确研究方向

尽管取得了一些学术成果，但投稿过程的反复失利与不确定性，让我陷入了一段巨大的迷茫。我开始激烈地自我怀疑：我真的适合这条孤独的道路吗？为了寻找答案，我做出了一个决定：主动走进业界，对我内心真正的热情进行一次彻底的“压力测试”。

我先加入了由前 IBM 中国区 AI 总架构师方杰先生带领的初创公司 KeyoneAI。在这里，我将前沿的生成式 AI 技术落地为产品，感受着快速迭代的脉搏，与团队写作与潜在用户交锋，做技术的布道者。

随后，我加入了世界教育者大会 (WWEC) 的组委会 [3]，作为唯一的技术负责人，我用技术赋能了一场上万人的国际化会议。在保障系统功能的同时促进团队数字化、开发内部工具支持 1000+ 复杂海报生成、链接 3000 观众、1000 展商、10+ 供应商和各种临时会议，每周 80 小时工作近 3 个月。尽管这些都带来了巨大的挑战和经历后的成就感，但我发现，它们都无法替代我在科研过程中，发现在某个技术方向上科研人员几年一突破、一年几突破，并且让自己参与到这种突破中时，所感受到的那种纯粹的、智识上的兴奋与心流。这段“刻意疏离”的探索，让我获得了前所未有的清晰：我最深的渴望，是对事物底层原理的追问，是对现有方案的驳斥与重建，是旁征博引、有效创新的构建。

更重要的是，这次探索帮助我将曾经困扰我的个人特质过度的自我审视与“内耗”，升华为我未来研究的独特视角。我比许多人更深刻地理解，**一个真正智能的系统，其标志不应是毫无瑕疵的单向推理，而恰恰是具备处理内在冲突、进行自我审视和迭代修正的能力**这正是人类“反思”与“纠结”的本质。

#### 5. 未来的研究蓝图：对齐语义，推理未来

现在，我带着更加清晰和坚定的目标来规划我的博士研究。我希望将我过去在**模型机制设计**上的经验，与对**人类认知过程**的思考结合起来，专注于探索大模型的两个核心能力：

1. **构建统一的语义空间**：我的首要目标是研究如何将各种信息，有效映射到一个统一的、可解析语义的隐空间中。这不止是在强调跨模态的映射，也是逻辑推理的基础。我相信，实现高效的跨模态语义对齐，是构建能够全面理解世界的基础模型的第一步。我在 QGFace 项目中处理混合质量数据的经验，为我探索如何对齐和融合异构信息提供了宝贵的实践基础。

2. **在隐空间中实现结构化的推理过程**：在实现语义对齐之后，我的进一步目标是在这个统一的隐空间内，设计并实现结构化的、类似“思维链”的推理路径。我希望模型不仅能得到答案，更能展现出一个可分解、可追溯的推理过程。这不仅能增强模型的可解释性和可靠性，也可能为解决更复杂的、需要多步逻辑的开放式问题提供新的途径。

总而言之，我的研究计划是双向并进的：一方面，通过**多模态对齐**让模型“看懂”更丰富的世界；另一方面，通过**潜在推理**让模型“想得”更清晰、更有逻辑。

#### 6. 结论

我相信，凭借我在设计模型内部机制方面的实践经验，以及对构建“语义对齐”与“潜在推理”相结合的 AI 系统的清晰规划，我已准备好迎接博士阶段的挑战。我渴望在一个富有创造力和支持性的环境中，为构建下一代更强大、更可信的人工智能系统贡献自己的力量。

## 参考文献

### References

- [1] Youzhe Song and Feng Wang. Coreface: Sample-guided contrastive regularization for deep face recognition. Pattern Recognition, 152:110483, 2024.
- [2] Youzhe Song and Feng Wang. Quality-guided joint training for mixed-quality face recognition. In 2024 IEEE International Conference on Automatic Face and Gesture Recognition (FG). IEEE, 2024.
- [3] Worldwide educators conference (wwec). <https://www.wwec820.com/>. Accessed: 2025-10-01.