**报告表格模型识别方案**

## **推荐模型：**

1.亚马逊模型

Amazon   Nova

2.谷歌模型

Gemini 2.0 Flash Lite

对于亚马逊的Nova模型

已经做的尝试：

1. 纯自然语言 依靠思维链 样本举例 等 小模型无法合理处理数据。会出现幻视的问题，误识别 错误数据；
2. 矩阵化处理 单轮对话获取结果。 小模型无法根据输入矩阵的规则，期间也会出现幻视问题；

基于以上尝试 无法对于Nova模型需要更细致的方案去处理 样本表格，目前希望对表格识别制定一套规则，并且通过多轮会话完成复杂的任务；

尝试的方式： 表格任务“分层模型+规则”设计思路、编码（霍夫曼编码）、降维等方法

你的问题总结为两个部分：

1. **目标**：想用“小模型”来降本增效，但小模型直接处理“多token任务”（如长文本、多字段表格）效果不好。
2. **思路探讨**：能否类比卷积神经网络的“分层特征提取”思想，针对表格任务，分几步（如先搞特征、后编规则等）搭出一套“以任务为导向”的小模型方案？

我用“白话+示意+分层讲解”来梳理——

## **一、类比CNN的多层特征提取思想做表格识别**

### **1. 卷积神经网络的分层思想回顾：**

* **层1**：低级特征（如边缘、线段）
* **层2**：中级特征（小物体、纹理）
* **层3**：高级特征（全局结构、语义）

### **2. 能否迁移到表格特征提取方案？**

可以！核心思路是，把大问题**拆细分层**，每层处理简单点儿，逐步编码更复杂的特征。  
大模型可以直接把一大坨输入吃完后理解，小模型则更适合组合式、流水线分层提取。

## **二、表格任务的“分层模型 + 规则”方案设计思路**

### **方案总览（分层+精细规则提升）：**

#### **第一层（低级特征抽取）**

* **目标**：提取单元格的基本内容、格式、字段类型等。
* **方式**：用非常小的模型+简单启发式，比如：
* 识别数字/金额/日期
* 基于正则、基础ML分类每个格子的类型

#### **第二层（结构关系识别）**

* **目标**：抽取格子之间的上下文关系，比如标题行、合并单元格、表格头部及内容区分。
* **方式**：用一点点更复杂的小模型，或针对结果组合一定规则（仿CNN方块滑窗提特征）。
* 例：根据相邻格子内容和位置推测表头
* 简单树模型、卡方过滤，或者模板引导

#### **第三层（语义/业务规则层）**

* **目标**：把抽取的基础和结构特征，组合为实用的“业务标签”或“任务决策”。
* **方式**：
* 多轮规则推理
* 简单的条件判定（如A格是金额且B格是“合计”则为总金额）
* 领域知识引导（可以Chat或人工把规则补上）

#### **第四层（端到端质量保障/纠错）**

* **目标**：发现模型可能遗漏或错判的case，通过规则或微调小模型进一步校正。
* **方式**：
* 结果自动对账/校验
* 引入人工review/异常提示
* 逐步用多轮对话把复杂情况“划分更多if-else”补全

## **三、可怎么组合“小模型”+“规则”做降本增效？**

* **模型部分**：每一层用最小化的模型，根据实际效果分步部署；难点调多点，小模型跑起来更快。
* **规则部分**：用人工/专家/多轮对话（甚至AI辅助）补充、修正模型不容易学到的细节或异常。
* **可持续优化**：统计推理难点分布，局部上加“专用小插件”或“小规则程序”，把典型难例单独处理。

**类比CNN**

* 卷积=基础特征提取
* 池化/跳跃连接=规则/上下文整合
* 全连接层=任务决策/最终输出
* Dropout/正则化=错误检测与纠错模块

## **四、举例**

**场景**：表格中自动找“发票金额”、“发票日期”、“合计金额”字段

1. 第1步（启发式/小模型）找出所有形如数值的单元格，初步筛“金额”
2. 第2步（模板规则/小模型）分析文本内容，判断“合计”、“日期”等关键词上下位置，组合特征关系
3. 第3步（专家规则/多轮对话/模型微调）处理易混淆case，引入细分判断
4. 第4步（异常检校/人工辅助）再抓漏网之鱼

## **五、结论（流程图式方案）**

1. **低级自动化 → 结构关系 → 任务语义 → 高级规则+人工协同**
2. 小模型只分担它能胜任的“自动特征”，靠规则/专家/经验补全理解深度
3. 类比CNN分层提特征、分层做决策，搭建降本增效流水线，逐步补齐准确率

如需具体的技术方案蓝图、代码范例或每层可选的小模型类型（比如轻量分类器、NER模型等），可以进一步告诉我详细任务内容，我可帮你细化到每一层的具体推荐！