

A Survey on Multimodal Large Language Models

```
graph LR; A[A Survey on Multimodal Large Language Models] --- B[M-IT (instruction tuning)]; A --- C[M-ICL (in context learning)]; A --- D[Multimodal Chain of Thought]; A --- E[LLM-Aided Visual Reasoning]; B --- B1[数据集的样式是有nstruction + response]; C --- C1[给出example，然后补全]; C --- C2[不更新参数，只是eval模型的类比推理能力？]; D --- D1[具体没看过论文，类似LLM的CoT];
```

M-IT (instruction tuning)

数据集的样式是有nstruction + response

M-ICL (in context learning)

给出example，然后补全

不更新参数，只是eval模型的类比推理能力？

Multimodal Chain of Thought

具体没看过论文，类似LLM的CoT

LLM-Aided Visual Reasoning